WEST Search History

Hide Items | Restore | Clear | Cancel

DATE: Monday, September 27, 2004

Hide?	Set Name	Query	Hit Count
		DB=PGPB,USPT,USOC,EPAB,JPAB,DWPI,TDBD; PLUR=YES; OP=AND	
☐	L1	pht.clm.	35
☐	L2	(polyhistidine near2 triad) or PhtX	24
☐	L3	(poly-histidine near2 triad)	0
☐	L4	(histidine near2 triad)	207
☐	L5	L4 same (strep or streptococcu$ or pneumoniae or pnumoni$)	8

END OF SEARCH HISTORY

Generate Collection   Print

Search Results - Record(s) 1 through 8 of 8 returned.

☐ 1. 20040081662. 08 Oct 03. 29 Apr 04. Vaccine. Hermand, Philippe, et al. 424/190.1; A61K039/02.

☐ 2. 20040052781. 14 Apr 03. 18 Mar 04. Vaccine compositions comprising Streptococcus pneumoniae polypeptides having selected structural motifs. Johnson, Leslie S., et al. 424/130.1; 424/185.1 435/100 A61K039/395 A61K039/00.

☐ 3. 20040005331. 13 Mar 03. 08 Jan 04. Vaccine compositions comprising Streptococcus pneumoniae polypeptides having selected structural motifs. Johnson, Leslie S., et al. 424/190.1; 530/350 536/23.7 A61K039/02 C07H021/04 C07K014/315.

☐ 4. 20040001836. 14 Apr 03. 01 Jan 04. Vaccine compositions comprising streptococcus pneumoniae polypeptides having selected structural motifs. Johnson, Leslie S., et al. 424/165.1; 424/190.1 A61K039/40 A61K039/02.

☐ 5. 20020197605. 18 Dec 00. 26 Dec 02. Novel Polynucleotides. Nakagawa, Satoshi, et al. 435/6; 435/287.2 435/91.2 C12Q001/68 C12P019/34 C12M001/34.

☐ 6. 6582706. 21 Dec 99; 24 Jun 03. Vaccine compositions comprising Streptococcus pneumoniae polypeptides having selected structural MOTIFS. Johnson; Leslie S., et al. 424/244.1; 424/184.1 424/185.1 424/190.1 424/237.1 435/320.1 435/69.1 530/350 536/23.1 536/23.7. A61K039/09.

☐ 7. WO 200222168A. New immunogenic composition for treating streptococcal infections in infants and elders, comprises two Streptococcus pneumoniae proteins selected from the poly histidine triad family and the choline binding protein family. HERMAND, P, et al. A61K000/00 A61K039/02 A61K039/09 A61K039/385 A61P011/00 A61P027/16 A61P031/04.

☐ 8. WO 200037105A. Vaccine useful for prophylaxis and treatment of pnemococcal infections such as otitis media, nasopharyngeal and bronchial infections, comprises Streptococcus pneumoniae proteins. ADAMOU, J E, et al. A61K038/00 A61K039/00 A61K039/02 A61K039/09 A61K039/395 A61K039/40 A61P031/04 A61P031/10 C07H021/04 C07K014/315 C07K014:315.

Generate Collection   Print

Terms	Documents
L4 same (strep or streptococcu$ or pneumoniae or pnumoni$)	8

Prev Page   Next Page   Go to Doc#

WEST Search History

Hide Items | Restore | Clear | Cancel
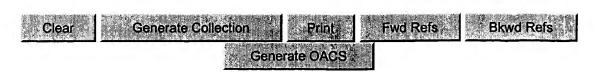
DATE: Monday, September 27, 2004

Hide?	Set Name	Query	Hit Count
		DB=USPT; PLUR=YES; OP=AND	
☐	L1	lxxc	5

END OF SEARCH HISTORY

Hit List

Search Results - Record(s) 1 through 5 of 5 returned.

☐ 1.   Document ID: US 6773880 B2

L1: Entry 1 of 5   File: USPT   Aug 10, 2004

DOCUMENT-IDENTIFIER: US 6773880 B2
TITLE: Streptococcus pneumoniale 37-kDa surface adhesion A protein

Detailed Description Text (73):
The complete nucleotide sequence on both strands of the 1.3-kb insert was
determined by cycle sequencing of the plasmid subclone using oligonucleotide
primers complementary to the sequence. These were made as sequence information
became available. The nucleotide sequence of the entire streptococcal insert is set
forth in the Sequence Listing as SEQ ID NO:1. The single open reading frame (ORF)
present, beginning at nucleotide (nt) 189 and ending at nt 1117, encodes the psaA
gene sequence. This ORF is 930 nt long and when amplified and subcloned into vector
systems such as pGEM (Promega, Madison, Wis.) and BAC-to-BAC.sup.tm expression
system (Bethesda Research Laboratories, Gaithersburg, Md.) expresses full-length
PsaA, reactive with anti-PsaA MAb antibodies. This ORF encodes a peptide of 309
amino acids with a deduced molecular weight of 34,598 and an isoelectric point of
5.23. Analysis of the peptide using the algorithm of Kyte and Doolittle (Kyte et
al. 1982. "A simple method for displaying the hydropathic character of a protein."
J. Mol. Biol. 157:105-132) shows that the peptide contains a major hydrophobic
region of 20 amino acids which encodes a putative leader sequence. This leader
contains the consensus sequence for signal peptidase cleavage (LXXC). Removal of
this leader would result in a peptide of molecular mass 32,465 with a predicted
isoelectric point of 4.97. A consensus sequence for a ribosomal binding site (Shine
et al. 1974. "The 3'-terminal sequence of E. coli 16S ribosomal RNA:
complementarity to nonsense triplets and ribosomal binding sites." Proc. Natl.
Acad. Sci. USA 71:1324-1346) is located 5 nt upstream of the ATG start codon.

Detailed Description Text (76):
A comparison of serotype 2 and 6B PsaAs shows almost complete identity: the
computed similarity value is 99.3. The eight base difference at the nucleotide
level translated into a difference at the peptide level of six amino acids with two
of the changes resulting in conservative substitutions. Further analyses and
comparisons of the serotype 6B sequence to the other five GenBank PsaA homologues
from viridans Streptococci and E. faecalis (Fenno et al. 1989. "Nucleotide sequence
analysis of a type 1 fimbrial gene of Streptococcus sanguis FW213." Infect. Immun.
57:3527-3533, Sampson et al. 1994. "Cloning and nucleotide sequence analysis of
psaA, the Streptococcus pneumoniae gene encoding a 37-kilodalton protein homologous
to previously reported Streptococcus sp. adhesins." Infect. Immun. 62:319-324,
Ganeshkumar et al. 1991. "Nucleotide sequence of a gene coding for a salvia-binding
protein (SsaB) from Streptococcus sanguis 12 and possible role of the protein in
coaggregation with actinomyces." Infect. Immun. 59:1093-1099, Kolenbrander et al.
1994. "Nucleotide sequence of the Streptococcus gordonii PK488 coaggregation
adhesin gene scaA and ATP-binding cassette." Infect. Immun. 62:4469-4480, and Lowe

All of the known Lral proteins are lipoproteins with the typical recognition sequence LxxC for the signal peptidase II at amino acid residue 16 or 17. The GBS homologue has a similar but slightly different sequence at this position (IAGC) with an exchange of leucine to isoleucine. LplA of Bacillus subtilis which has been shown to be a lipoprotein by radiolabelling with palmitate, is another example of this atypical recognition sequence in a gram-positive organism (Sutcliffe and Russel 1995).

| Full | Title | Citation | Front | Review | Classification | Date | Reference | Sequences | Attachments | Claims | KWIC | Draw D |

4. Document ID: US 6217884 B1

L1: Entry 4 of 5   File: USPT   Apr 17, 2001

DOCUMENT-IDENTIFIER: US 6217884 B1
TITLE: Streptococcus pneumoniae 37-kDa surface adhesin a protein

Detailed Description Text (66):
The complete nucleotide sequence on both strands of the 1.3-kb insert was determined by cycle sequencing of the plasmid subclone using oligonucleotide primers complementary to the sequence. These were made as sequence information became available. The nucleotide sequence of the entire streptococcal insert is set forth in the Sequence Listing as SEQ ID NO:1. The single open reading frame (ORF) present, beginning at nucleotide (nt) 189 and ending at nt 1117, encodes the psaA gene sequence. This ORF is 930 nt long and when amplified and subcloned into vector systems such as pGEM (Promega, Madison, Wis.) and BAC-to-BA.TM. expression system (Bethesda Research Laboratories, Gaithersburg, Md.) expresses full-length PsaA, reactive with anti-PsaA MAb antibodies. This ORF encodes a peptide of 309 amino acids with a deduced molecular weight of 34,598 and an isoelectric point of 5.23. Analysis of the peptide using the algorithm of Kyte and Doolittle (Kyte et al. 1982. "A simple method for displaying the hydropathic character of a protein." J. Mol. Biol. 157:105-132) shows that the peptide contains a major hydrophobic region of 20 amino acids which encodes a putative leader sequence. This leader contains the consensus sequence for signal peptidase cleavage (LXXC). Removal of this leader would result in a peptide of molecular mass 32,465 with a predicted isoelectric point of 4.97. A consensus sequence for a ribosomal binding site (Shine et al. 1974. "The 3'-terminal sequence of E. coli 16S ribosomal RNA: complementarity to nonsense triplets and ribosomal binding sites." Proc. Natl. Acad. Sci. USA 71:1324-1346) is located 5 nt upstream of the ATG start codon.

Detailed Description Text (69):
A comparison of serotype 2 and 6B PsaAs shows almost complete identity: the computed similarity value is 99.3. The eight base difference at the nucleotide level translated into a difference at the peptide level of six amino acids with two of the changes resulting in conservative substitutions. Further analyses and comparisons of the serotype 6B sequence to the other five GenBank PsaA homologues from viridans Streptococci and E. faecalis (Fenno et al. 1989. "Nucleotide sequence analysis of a type 1 fimbrial gene of Streptococcus sanguis FW213." Infect. Immun. 57:3527-3533, Sampson et al. 1994. "Cloning and nucleotide sequence analysis of psaA, the Streptococcus pneumoniae gene encoding a 37-kilodalton protein homologous to previously reported Streptococcus sp. adhesins." Infect. Immun. 62:319-324, Ganeshkumar et al. 1991. "Nucleotide sequence of a gene coding for a salvia-binding protein (SsaB) from Streptococcus sanguis 12 and possible role of the protein in coaggregation with actinomyces." Infect. Immun. 59:1093-1099, Kolenbrander et al. 1994. "Nucleotide sequence of the Streptococcus gordonii PK488 coaggregation

adhesin gene scaA and ATP-binding cassette." Infect. Immun. 62:4469-4480, and Lowe
et al. 1995. "Cloning of an Enterococcus faecalis endocarditis antigen: homology
with some adhesins from oral streptococci." Infect. Immun 63:703-706) revealed
significant sequence similarity between them. Sequence identities were 81%, 81%,
77%, 82%, and 57%, respectively, for PsaA (S. pneumoniae strain R36A), SsaB (S.
sanguis), FimA (S. parasanguis), ScaA (S. gordonii) and EfaA (E. faecalis).
Additionally, all six sequences showed great similarity in organization. They have a
a hydrophobic leader peptide containing the prolipoprotein consensus sequence LXXC
(for signal peptidase II cleavage) within the first 17-20 amino acids. This N-
terminal leader sequence appears to represent the area of greatest variability. It
is followed by a region of high similarity from amino acids 36-150. The region from
150 to 198 is a variable region and is followed by another conserved region (198-
309).

| Full | Title | Citation | Front | Review | Classification | Date | Reference | Sequences | Attachments | Claims | KWIC | Draw. D |

☐ 5.   Document ID: US 5854416 A

L1: Entry 5 of 5   File: USPT   Dec 29, 1998

DOCUMENT-IDENTIFIER: US 5854416 A
TITLE: Streptococcus pneumoniae 37-KDA surface adhesin a protein and nucleic acids
coding therefor

Detailed Description Text (76):
The complete nucleotide sequence on both strands of the 1.3-kb insert was
determined by cycle sequencing of the plasmid subclone using oligonucleotide
primers complementary to the sequence. These were made as sequence information
became available. The nucleotide sequence of the entire streptococcal insert is set
forth in the Sequence Listing as SEQ ID NO:1. The single open reading frame (ORF)
present, beginning at nucleotide (nt) 189 and ending at nt 1117, encodes the psaA
gene sequence. This ORF is 930 nt long and when amplified and subcloned into vector
systems such as pGEM (Promega, Madison, Wis.) and BAC-to-BAC.TM. expression system
(Bethesda Research Laboratories, Gaithersburg, Md.) expresses full-length PsaA,
reactive with anti-PsaA MAb antibodies. This ORF encodes a peptide of 309 amino
acids with a deduced molecular weight of 34,598 and an isoelectric point of 5.23.
Analysis of the peptide using the algorithm of Kyte and Doolittle (Kyte et al.
1982. "A simple method for displaying the hydropathic character of a protein." J.
Mol. Biol. 157:105-132) shows that the peptide contains a major hydrophobic region
of 20 amino acids which encodes a putative leader sequence. This leader contains
the consensus sequence for signal peptidase cleavage (LXXC). Removal of this leader
would result in a peptide of molecular mass 32,465 with a predicted isoelectric
point of 4.97. A consensus sequence for a ribosomal binding site (Shine et al.
1974. "The 3'-terminal sequence of E. coil 16S ribosomal RNA: complementarity to
nonsense triplets and ribosomal binding sites." Proc. Natl. Acad. Sci. U.S.A.
71:1324-1346) is located 5 nt upstream of the ATG start codon.

Detailed Description Text (79):
A comparison of serotype 2 and 6B PsaAs shows almost complete identity: the
computed similarity value is 99.3. The eight base difference at the nucleotide
level translated into a difference at the peptide level of six amino acids with two
of the changes resulting in conservative substitutions. Further analyses and
comparisons of the serotype 6B sequence to the other five GenBank PsaA homologues
from viridans Streptococci and E. faecalis (Fenno et al. 1989. "Nucleotide sequence
analysis of a type 1 fimbrial gene of Streptococcus sanguis FW213." Infect. Immun.

57:3527-3533, Sampson et al. 1994. "Cloning and nucleotide sequence analysis of psaA, the Streptococcus pneumoniae gene encoding a 37-kilodalton protein homologous to previously reported Streptococcus sp. adhesins." Infect. Immun. 62:319-324, Ganeshkumar et al. 1991. "Nucleotide sequence of a gene coding for a salvia-binding protein (SsaB) from Streptococcus sanguis 12 and possible role of the protein in coaggregation with actinomyces." Infect. Immun. 59:1093-1099, Kolenbrander et al. 1994. "Nucleotide sequence of the Streptococcus gordonii PK488 coaggregation adhesin adhesin gene scaA and ATP-binding cassette." Infect. Immun. 62:4469-4480, and Lowe et al. 1995. "Cloning of an Enterococcus faecalis endocarditis antigen: homology with some adhesins from oral streptococci. " Infect. Immun 63:703-706) revealed significant sequence similarity between them. Sequence identities were 81%, 81%, 77%, 82%, and 57%, respectively, for PsaA (S. pneumoniae strain R36A), SsaB (S. sanguis), FimA (S. parasanguis), ScaA (S. gordonii) and EfaA (E. faecalis). Additionally, all six sequences showed great similarity in organization. They have a a hydrophobic leader peptide containing the prolipoprotein consensus sequence LXXC (for signal peptidase II cleavage) within the first 17-20 amino acids. This N-terminal leader sequence appears to represent the area of greatest variability. It is followed by a region of high similarity from amino acids 36-150. The region from 150 to 198 is a variable region and is followed by another conserved region (198-309).

| Full | Title | Citation | Front | Review | Classification | Date | Reference | Sequences | Attachments | Claims | KWIC | Draw De |

| Clear | Generate Collection | Print | Fwd Refs | Bkwd Refs |

Generate OACS

Terms	Documents
lxxc	5

Display Format: KWIC   Change Format

Previous Page   Next Page   Go to Doc#

First Hit   Fwd Refs

    L1: Entry 1 of 5   File: USPT   Aug 10, 2004

US-PAT-NO: 6773880
DOCUMENT-IDENTIFIER: US 6773880 B2

TITLE: Streptococcus pneumoniale 37-kDa surface adhesion A protein

DATE-ISSUED: August 10, 2004

INVENTOR-INFORMATION:
NAME	CITY	STATE	ZIP CODE	COUNTRY
Sampson; Jacquelyn	College Park	GA		
Russell; Harold	Efland	NC		
Tharpe; Jean A.	Lithonia	GA		
Ades; Edwin W.	Atlanta	GA		
Carlone; George M.	Stone Mountain	GA		

ASSIGNEE-INFORMATION:
NAME	CITY	STATE	ZIP CODE	COUNTRY	TYPE CODE
The United States of America as represented by the Department of Health and Human Services	Washington	DC			06

APPL-NO: 09/ 754809   [PALM]
DATE FILED: January 3, 2001

PARENT-CASE:
This application is a divisional of and claims the benefit of priority from, U.S.
Ser. No. 09/221,753, filed Dec. 28, 1998 now U.S. Pat. No. 6,217,884, which is a
divisional of U.S. Ser. No. 08/715,131, filed Sep. 17, 1996, now U.S. Pat. No.
5,854,416, which is a continuation-in-part of U.S. Ser. No. 08/222,179, filed Apr.
4, 1994, now abandoned which is a continuation-in-part of U.S. Ser. No. 07/791,377,
filed Sep. 17, 1991, now U.S. Pat. No. 5,422,427, which applications are hereby
incorporated herein by reference.

INT-CL: [07] C12 Q 1/68

US-CL-ISSUED: 435/6; 536/23.7, 536/24.32, 536/24.33
US-CL-CURRENT: 435/6; 536/23.7, 536/24.32, 536/24.33

FIELD-OF-SEARCH: 435/6, 536/23.7, 536/24.32, 536/24.33

PRIOR-ART-DISCLOSED:

                        U.S. PATENT DOCUMENTS

              Search Selected   Search ALL   Clear

	PAT-NO	ISSUE-DATE	PATENTEE-NAME	US-CL
☐	4762713	August 1988	Anderson	
☐	4789735	December 1988	Frank et al.	
☐	4894362	January 1990	Yamaguchi et al.	
☐	5037760	August 1991	Smith et al.	
☐	5130417	July 1992	Stanley et al.	
☐	5422427	June 1995	Russell et al.	
☐	5854416	December 1998	Sampson et al.	

FOREIGN PATENT DOCUMENTS

FOREIGN-PAT-NO	PUBN-DATE	COUNTRY	US-CL
0206 852	December 1986	EP	
0 429 816	June 1991	EP	

OTHER PUBLICATIONS

Sampson, et al., "Immonulogic Characterization of a Monoclonal Antibody to Streptococcus Pneumoniae Pneumococcal Surface Adhesin A (PsaA) Protein," Centers for for Disease Control & Prevention, Atlanta, Ga 30333.

Novak, et al., "Penicillin tolerance genes of Streptococcus pneumoniae: the ABC-type type manganese permease complex Psa," Molecular Microbiology, 29:1285-1296 (1998).

Lawrence, et al., "The crystal structure of pneumococcal surface antigen PsaA reveals a metal-binding site and a novel structure for a putative ABC-type binding protein," Structure, 6:1553-1561, (Dec. 1998).

Amit, et al., "Three-Dimensional Structure of a Antigen-Antibody Complex at 2.8 Angstrom Resolution," Science, 233:747-753 (Apr. 1986).

Kolenbrander, et al., "Nucleotide Sequence of the Streptococcus gordonii PK488 Coaggregation Adhesin Gene, scaA, and ATP-Binding Cassette," Infection and Immunity, Immunity, 62(10):4469-4480, (Oct. 1994).

Talkington, et al., "A 43-Kilodalton Pneumococcal Surface Protein, PspA: Isolation, Protective Abilities, and Structural Analysis of the Amino-Terminal Sequence," Infection and Immunity, 59(4):1285-1289, (Apr. 1991).

Sampson, et al., "Molecular Cloning of the Gene Encoding the 37-Kilodalton Protein of Streptococcus pneumoniae," Abstracts of the 91.sup.st General Meeting of the Am. Society for Microbiology Annual Meeting; p. 97, (1991).

Russell, et al., "Assay for Antibodies Against a Species-Specific Streptococcus Pneumoniae Antigen in Patients with Pneumococcal Disease," Abstracts of the 91.sup.st General Meeting of the Am. Society for Microbiology; p. 434 (1991).

Russell, et al., Isolation and Purification of a Species-Specific Streptococcus pneumoniae Protein Antigen by Isoelectric focusing, Abstracts of the Annual Meeting- Meeting--1990 of the Am. Society for Microbiology Annual Meeting; p. 436 (1990).

Ganeshkumar, et al., "Nucleotide Sequence of a Gene Coding for a Saliva-Binding Protein (SsaB) from Streptococcus sanguis 12 and Possible Role of the Protein in Coaggregation with Actinomyces," Infection and Immunity, 59(3):1093-1099, (Mar. 1991).

Russell, et al., "Investigation of Streptococcus Pneumoniae Components for Immunodiagnostic Markers," Abstracts of the 89.sup.th Annual Meeting of the Am. Society for Microbiology; p. 489 (1989).

Van de Wijgert, et al., "Immunogenicity of Streptococcus Pneumoniae Type 14 Capsular Capsular Polysaccharide: Influence of Carriers and Adjuvants on Isotype Distribution," Infection & Immunity, 59(8):2750-2757 (Aug. 1991).

Audibert and Lise, "Adjuvants: Current Status, Clinical Perspectives and Future Prospects," Immunology Today, 14(6):281-284 (1993).

Vella, et al., "Immunogenicity of Conjugate Vaccines Consisting of Pneumococcal Capsular Polysaccharide Types . . . ," Infection and Immunity, 60(12):4977-4983 (Dec. 1992).

Green, et al., "Evaluation of Mixtures of Purified Haemophilus Influenzae Outer Membrane Proteins . . . " Infection and Immunity, 61(5):1950-1957 (May 1993).

Bixler and Atassi, "B Cell Recognition of Protein Antigens--Perspectives From The Submolecular Level." In: Synthetic Vaccines, (Arnon, ed.) 1:39--71, CRC Press, Inc., Inc., Boca Raton, FL. (1987).

Berry and Paton, "Sequence Heterogeneity of PsaA, a 37-Kilodalton Putative Adhesin Essential for Virulence of Streptococcus pneumoniae," Infection and Immunity, 64(12):5255-5262 (Dec. 1996).

Russell, et al., "Monoclonal Antibody Recognizing a Species-Specific Protein from Streptococcus pneumoniae," J. Clin. Microbiol. 28:2191-2195, (Oct. 1990).

Briles, D.E., et al., "Antipneumococcal Effects of C-Reactive Protein and Monoclonal Monoclonal Antibodies to Pneumococcal Cell Wall and Capsular Antigens," Infection and Immunity, 57(5):1457-1464. (May 1989).

Fenno, et al., "Nucleotide Sequence Analysis of a Type 1 Fimbral Gene of Streptococcus sanguis FW213," Infection and Immunity, 57(11):3527-3533, (Nov. 1989). 1989).

Advertisement offering custom DNA/Peptides; total gene synthesis. Science, 240:362, (1988).

Fives-Taylor, et al., "Expression of Streptococcus sanguis Antigens in Escherichia Coli: Cloning of a Structural Gene for Adhesion Fimbriae," Infection and Immunity, 55(1):123-128, (Jan. 1987).

Clark-Lewis, et al., "Automated Chemical Synthesis of a Protein Growth Factor for Hemopoietic Cells, Interleukin-3," Science, 231:134-139, (Jan. 1986).

Caruthers, et al., "Gene Synthesis Machines: DNA Chemistry and Its Uses," Science, 230:281-285, (Oct. 1985).

Tharpe and Russell, "Purification and Seroreactivity of Pneumococcal Surface Adhesin Adhesin A (PsaA)," Clinical and Diagnostic Laboratory Immunology, 3(2):227-229, (Mar. 1996).

Sampson, et al., "Cloning and Nucleotide Sequence Analysis of psaA, the Streptococcus pneumoniae Gene Encoding a 37-Kilodalton Protein Homologous to Previously Reported Streptococcus sp. Adhesins," Infection and Immunity, 62(1):319-324, (Jan. 1994).

Sampson, et al., "Conservation of the 37-kDa Protein Gene among Streptococcus pneumoniae Serotypes," Interscience Conference of Antimicrobial Agents and Chemotherapy (ICAAC), (Sep. 17, 1995).

Yother and Briles, "Structural Properties and Evolutionary Relationships of PspA, a Surface Protein of Streptococcus pneumoniae, as Revealed by Sequence Analysis," J. of Bacteriology, 174(2):601-609, (Jan. 1992).

Waltman, et al., "Variation in the molecular weight of PspA (pneumococcal surface of of protein A) among Streptococcus pneumoniae," Microbial Pathogenesis, 8(1):61-69 (Jan. 1990).

Talkington, et al., "The 37-kDa Protein of Streptococcus pneumoniae Protects Mice against Fatal Challenge," Abstracts of the General Meeting of the American Society for Microbiology, New Orleans, Louisiana, cover page and E-29 on p. 149 (1992).

McDaniel, L.S., et al., "PspA, a Surface Protein of Streptococcus pneumoniae, Is Capable of Eliciting Protection against Pneumococci of More Than One Capsular Type," Type," Infection and Immunity, 59(1):222-228, (Jan. 1991).

McDaniel, L.S., et al., "Use of Insertional Inactivation To Facilitate Studies of Biological Properties of Pneumococcal Surface Protein A (PspA)," J. Experimental Medicine, 165:381-394, (Feb. 1987).

ART-UNIT: 1645

PRIMARY-EXAMINER: Graser; Jennifer E.

ATTY-AGENT-FIRM: Needle & Rosenberg, P.C.

ABSTRACT:

The invention provides a nucleic acid encoding the 37-kDa protein from Streptococcus Streptococcus pneumoniae. Also provided are isolated nucleic acids comprising a unique fragment of at least 10 nucleotides of the 37-kDa protein. The invention also also provides purified polypeptides encoded by the nucleic acid encoding the 37-kDa protein from and the nucleic acids comprising a unique fragment of at least 10 nucleotides of the 37-kDa protein. Also provided are antibodies which selectively binds the polypeptides encoded by the nucleic acid encoding the 37-kDa protein and the nucleic acids comprising a unique fragment of at least 10 nucleotides of the 37-37-kDa protein. Also provided are vaccines comprising immunogenic polypeptides encoded by the nucleic acid encoding the 37-kDa protein and the nucleic acids comprising a unique fragment of at least 10 nucleotides of the 37-kDa protein. Further provided is a method of detecting the presence of Streptococcus pneumoniae in a sample comprising the steps of contacting a sample suspected of containing Streptococcus pneumoniae with nucleic acid primers capable of hybridizing to a nucleic acid comprising a portion of the nucleic acid encoding the 37-kDa protein, amplifying the nucleic acid and detecting the presence of an amplification product, the presence of the amplification product indicating the presence of Streptococcus pneumoniae in the sample. Further provided are methods of detecting the presence of Streptococcus pneumoniae in a sample using antibodies or antigens, methods of preventing and treating Streptococcus pneumoniae infection in a subject.

1 Claims, 0 Drawing figures

Generate Collection | Print

Search Results - Record(s) 1 through 5 of 5 returned.

☐ 1. 6773880. 03 Jan 01; 10 Aug 04. Streptococcus pneumoniale 37-kDa surface adhesion A protein. Sampson; Jacquelyn, et al. 435/6; 536/23.7 536/24.32 536/24.33. C12Q001/68.

☐ 2. 6582706. 21 Dec 99; 24 Jun 03. Vaccine compositions comprising Streptococcus pneumoniae polypeptides having selected structural MOTIFS. Johnson; Leslie S., et al. 424/244.1; 424/184.1 424/185.1 424/190.1 424/237.1 435/320.1 435/69.1 530/350 536/23.1 536/23.7. A61K039/09.

☐ 3. 6406883. 25 Sep 98; 18 Jun 02. Lmb gene of Streptococcus agalactiae. Lutticken; Rudolf, et al. 435/69.1; 424/244.1 435/243 435/252.3 435/253.4 435/320.1 435/69.3 536/23.7. C12P021/06.

☐ 4. 6217884. 28 Dec 98; 17 Apr 01. Streptococcus pneumoniae 37-kDa surface adhesin a protein. Sampson; Jacquelyn S., et al. 424/244.1; 424/184.1 424/190.1 424/200.1 435/69.1 435/69.3 435/71.1 530/350 536/23.7. A61K039/09.

☐ 5. 5854416. 17 Sep 96; 29 Dec 98. Streptococcus pneumoniae 37-KDA surface adhesin a protein and nucleic acids coding therefor. Sampson; Jacquelyn S., et al. 536/23.7; 424/244.1 435/320.1 536/23.1. C07H021/04.

Generate Collection | Print

Terms	Documents
lxxc	5

Prev Page   Next Page   Go to Doc#

Generate Collection   Print

L1: Entry 5 of 35   File: PGPB   Apr 29, 2004

PGPUB-DOCUMENT-NUMBER: 20040081662
PGPUB-FILING-TYPE: new
DOCUMENT-IDENTIFIER: US 20040081662 A1

TITLE: Vaccine

PUBLICATION-DATE: April 29, 2004

INVENTOR-INFORMATION:

NAME	CITY	STATE	COUNTRY	RULE-47
Hermand, Philippe	Rixensart		BE	
Laferriere, Craig A.J.	Rixensart		BE	
Lobet, Yves	Rixensart		BE	
Poolman, Jan	Rixensart		BE	

US-CL-CURRENT: 424/190.1

CLAIMS:

1. An immunogenic composition comprising at least 2 S. pnumoniae proteins wherein one of the proteins is selected from the polyhistidine triad family (PhtX) and another protein is selected from the group consisting of Choline Binding Protein family (CbpX), CbpX truncates, LyX family, LytX truncates, CbpX truncate-LytX truncate chimeric proteins, pneumolysin (Ply), PspA, PsaA, Sp128, Sp101, Sp130, Sp125 and Sp133.

2. An immunogenic composition comprising at least 2 S. pneumoniae proteins wherein one of the proteins is selected from the group consisting of Choline Binding Protein family (CbpX), CbpX truncates, and CbpX truncate-LytX truncate chimeric proteins and another protein selected from the group consisting of polyhistidine triad family (PhtX), LytC, pneumolysin (Ply), PsaA, and Sp128.

3. The immunogenic composition of claims 1-2 wherein Pht is PhtA, PhtB or PhtD.

4. The immunogenic composition of claims 1-3 wherein CbpX is CbpA or PspC.

5. The immunogenic composition of claims 1-4 additionally comprising an adjuvant.

6. A vaccine comprising the immmogenic composition of claim 5.

7. A method of eliciting an immune response by immunising a mammal with the immunogenic composition of claims 1-5.

8. A method of preventing or ameliorating Streptococcus infection in patients over 55 years of age, comprising administering a safe and effective amount of the vaccine of claim 6 to said patients.

9. Use of the vaccine of claim 6 in the manufacture of a medicament for prevention of pneumonia in patients over 55 years of age.

10. A method of preventing or ameliorating Otitis media in infants, comprising administering a safe and effective amount of the vaccine of claim 6 to said patients.

11. A method of making a vaccine as claimed in claim 6 comprising the steps of: selecting and isolating two different S. pneumonia proteins; and mixing said proteins together with a pharmaceutically acceptable carrier.

<u>Previous Doc</u>   <u>Next Doc</u>   <u>Go to Doc#</u>

☐ Generate Collection   Print

L1: Entry 5 of 35   File: PGPB   Apr 29, 2004

PGPUB-DOCUMENT-NUMBER: 20040081662
PGPUB-FILING-TYPE: new
DOCUMENT-IDENTIFIER: US 20040081662 A1

TITLE: Vaccine

PUBLICATION-DATE: April 29, 2004

INVENTOR-INFORMATION:

NAME	CITY	STATE	COUNTRY	RULE-47
Hermand, Philippe	Rixensart		BE	
Laferriere, Craig A.J.	Rixensart		BE	
Lobet, Yves	Rixensart		BE	
Poolman, Jan	Rixensart		BE	

APPL-NO: 10/ 380563   [PALM]
DATE FILED: October 8, 2003

FOREIGN-APPL-PRIORITY-DATA:

COUNTRY	APPL-NO	DOC-ID	APPL-DATE
GB	00227421	2000GB-00227421	September 15, 2000

PCT-DATA:

DATE-FILED	APPL-NO	PUB-NO	PUB-DATE	371-DATE	102(E)-DATE
Sep 12, 2001	PCT/EP01/10570				

INT-CL: [07] A61 K 39/02

US-CL-PUBLISHED: 424/190.1
US-CL-CURRENT: 424/190.1

REPRESENTATIVE-FIGURES: NONE

ABSTRACT:

The present invention relates to a combination of 2 or more S pneumoniae proteins,
their manufacture and use in medicine as a vaccine. Such combinations are
particularly useful for the protection of infants and elderly against streptococcal
infection.

tr Q97QM9 **Conserved domain protein [SP1174] [Streptococcus pneumoniae]** 819 AA

Score = 423 bits (1087), Expect = e-116
Identities = 211/357 (59%), Positives = 271/357 (75%), Gaps = 11/357 (3%)

```
Query: 1    MKFSKKYIAAGSAVIVSLSLCAYALNQHRS-QENKDNNRVSYVDGSQSSQKSENLTPDQV 59
            MK +KKY+A GS  +++LS+C+Y L ++++ Q+ K++NRV+Y+DG Q+ QK+ENLTPD+V
Sbjct: 1    MKINKKYLA-GSVAVLALSVCSYELGRYQAGQDKKESNRVAYIDGDQAGQKAENLTPDEV 59

Query: 60   SQKEGIQAEQIVIKITDQGYVTSHGDHYHYYNGKVPYDALFSEELLMKDPNYQLKDADIV 119
            S++EGI AEQIVIKITDQGYVTSHGDHYHYYNGKVPYDA+ SEELLMKDPNYQLKD+DIV
Sbjct: 60   SKREGINAEQIVIKITDQGYVTSHGDHYHYYNGKVPYDAIISEELLMKDPNYQLKDSDIV 119

Query: 120  NEVKGGYIIKVDGKYYVYLKDAAHADNVRTKDEINRQKQEHVKD-NEKVNSNVAVARSQG 178
            NE+KGGY+IKV+GKYYVYLKDAAHADN+RTK+EI RQKQE  + N + ++ VA AR+QG
Sbjct: 120  NEIKGGYVIKVNGKYYVYLKDAAHADNIRTKEEIKRQKQERSHNHNSRADNAVAAARAQG 179

Query: 179  RYTTNDGYVFNPADIIEDTGNAYIVPHGGHYHYIPXXXXXXXXXXXXXXXXXXXXNMQPSQ 238
            RYTT+DGY+FN +DIIEDTG+AYIVPHG HYHYIP                    Q S+
Sbjct: 180  RYTTDDGYIFNASDIIEDTGDAYIVPHGDHYHYIP--KNELSASELAAAEAYWNGKQGSR 237

Query: 239  LSYSSTASDNNTQ---SVAKGSTSKPA---NKSENLQSLLKELYDSPSAQRYSESDGLVF 292
             S SS+ + N Q    S      T P    N+ EN+ SLL+ELY  P ++R+ ESDGL+F
Sbjct: 238  PSSSSSYNANPAQPRLSENHNLTVTPTYHQNQGENISSLLRELYAKPLSERHVESDGLIF 297

Query: 293  DPAKIISRTPNGVAIPHGDHYHFIPYSKLSALEEKIARRVPISGTGSTVSTNAKPNE 349
            DPA+I SRT  GVA+PHG+HYHFIPY ++S LE++IAR +P+   +   +++P E
Sbjct: 298  DPAQITSRTARGVAVPHGNHYHFIPYEQMSELEKRIARIIPLRYRSNHWVPDSRPEE 354
```

	ExPASy Home page	**Site Map**	**Search ExPASy**	**Contact us**	**Proteomics tools**	**Swiss-Prot**

Search [Swiss-Prot/TrEMBL ▼] for [bvh] [Go] [Clear]

```
===============================================================================

                    Welcome to the SIB BLAST Network Service

If results of this search are reported or published, please mention that
the computation was performed at the SIB using the BLAST network service.
The SIB BLAST network service uses a server developed at SIB and the NCBI
BLAST 2 software.

In case of problems, please read the online BLAST help.
If your question is not covered, please contact <helpdesk@expasy.org>.

NCBI BLAST program reference [PMID:9254694]:
Altschul S.F., Madden T.L., Schäffer A.A., Zhang J., Zhang Z., Miller W.,
Lipman D.J.  Gapped BLAST and PSI-BLAST: a new generation of protein
database search programs.  Nucleic Acids Res. 25:3389-3402(1997).


===============================================================================

Query length: 1039 AA (of which 3% low-complexity regions filtered out)
Date run: 2004-09-27 14:41:04 UTC+0100 on sib-gm1.unil.ch
Program: NCBI BLASTP 1.5.4-Paracel [2003-06-05]
Database: EXPASY/UniProt
          1,569,420 sequences; 502,311,625 total letters
```

Taxonomic view	NiceBlast view	Printable view

List of potentially matching sequences

Send selected sequences to [Clustal W (multiple alignment) ▼] [Submit Query]
[Select up to...]

☐ Include query sequence

	Db	AC	Description	Score	E-val
☐	tr	Q8DQ07	Pneumococcal histidine triad protein E [phtE] [Strepto...	2017	(
☐	tr	Q6WNQ7	Surface protein BVH-3 [bvh-3] [Streptococcus pneumoniae]	2017	(
☐	tr	Q9ANY1	Pneumococcal histidine triad protein E precursor (Hypo...	2006	(
☐	tr	Q6WNQ5	Surface protein BVH-3 (Fragment) [bvh-3] [Streptococcu...	1968	(
☐	tr	Q8CWR4	Histidine Motif-Containing protein [phpA] [Streptococc...	442	e-1
☐	tr	Q8DPQ2	Pneumococcal histidine triad protein A [phtA] [Strepto...	437	e-1
☐	tr	Q9AG74	PhpA [phpA] [Streptococcus pneumoniae]	437	e-1
☐	tr	Q9AHT9	Pneumococcal histidine triad A protein [phtA] [Strepto...	435	e-1
☐	tr	Q8DQ08	Pneumococcal histidine triad protein D [phtD] [Strepto...	434	e-1
☐	tr	Q6T8D7	Pneumococcal protein BVH-11-3 [Streptococcus pneumoniae]	431	e-1

☐	tr	Q97QM8	Conserved domain protein [SP1175] [Streptococcus pneum...	424	e-]
☐	tr	Q9ANY2	Pneumococcal histidine triad protein D precursor (Frag...	423	e-]
☐	tr	Q97QM9	Conserved domain protein [SP1174] [Streptococcus pneum...	423	e-]
☐	tr	Q9ANY3	Pneumococcal histidine triad protein B precursor (Frag...	423	e-]
☐	tr	Q6WNQ3	Surface protein BVH-11 (Fragment) [bvh-11] [Streptococ...	417	e-]
☐	tr	Q6WNP8	Surface protein BVH-11-2 (Fragment) [bvh-11-2] [Strept...	417	e-]
☐	tr	Q6WNQ1	Surface protein BVH-11 (Fragment) [bvh-11] [Streptococ...	415	e-]
☐	tr	Q6WNP5	Surface protein BVH-11-2 (Fragment) [bvh-11-2] [Strept...	415	e-]
☐	tr	Q6WNP9	Surface protein BVH-11-2 (Fragment) [Streptococcus pne...	410	e-]
☐	tr	Q6T304	Surface protein BVH-11 (Fragment) [bvh11] [Streptococc...	410	e-]
☐	tr	Q6WNQ0	Surface protein BVH-11 (Fragment) [bvh-11] [Streptococ...	405	e-]
☐	tr	Q6WNP6	Surface protein BVH-11-2 (Fragment) [bvh-11-2] [Strept...	400	e-]
☐	tr	Q8NZ82	Hypothetical protein spyM18_2072 [spyM18_2072] [Strept...	306	1e-
☐	tr	Q8E4U1	Hypothetical protein gbs1306 [gbs1306] [Streptococcus ...	305	2e-
☐	tr	Q8DZ81	Streptococcal histidine triad family protein [SAG1233]...	305	2e-
☐	tr	Q9ZHG7	Hypothetical protein [Streptococcus agalactiae]	305	3e-
☐	tr	Q99XV4	Hypothetical protein SPy2006 [SPy2006] [Streptococcus ...	304	5e-
☐	tr	Q8K5Q1	Histidine triad protein [SpyM3_1724] [Streptococcus py...	304	7e-
☐	tr	Q93GT5	Histidine triad protein of group A streptococci [htpA]...	302	2e-
☐	tr	Q8E338	Hypothetical protein gbs1925 [gbs1925] [Streptococcus ...	301	4e-
☐	tr	Q877Y2	Hypothetical protein SPs1722 [SPs1722] [Streptococcus ...	291	3e-
☐	tr	Q9AE21	Hypothetical protein (Fragment) [Streptococcus agalact...	241	7e-
☐	tr	Q8DQ06	Pneumococcal histidine triad protein E, truncation [ph...	121	6e-
☐	tr	Q8E029	Hypothetical protein SAG0907 [SAG0907] [Streptococcus ...	111	8e-
☐	tr	Q8E5R2	Hypothetical protein gbs0918 [gbs0918] [Streptococcus ...	109	2e-
☐	tr	Q8P0G5	Putative internalin A [spyM18_1373] [Streptococcus pyo...	99	3e-
☐	tr	Q8K714	Putative internalin A [inlA] [Streptococcus pyogenes (...	97	1e-
☐	tr	Q79XH7	Putative internalin A [SPs0825] [Streptococcus pyogene...	97	1e-
☐	tr	Q99Z76	Putative internalin A [inlA] [Streptococcus pyogenes]	92	7e-
☐	tr	Q6HCJ0	Cell division protein [ftsK] [Bacillus thuringiensis (...	47	0.(
☐	sp	P45386	IGA4_HAEIN Immunoglobulin A1 protease precursor (EC 3....	46	0.(
☐	tr	Q8MNS0	Hypothetical protein F12F3.2 [F12F3.2] [Caenorhabditis...	46	0.(
☐	tr	Q8ISF7	2MDa_1 protein [isof] [Caenorhabditis elegans]	46	0.(
☐	tr	Q8ISF6	2MDa_2 protein [isof] [Caenorhabditis elegans]	46	0.(
☐	sp	P16053	NFM_CHICK Neurofilament triplet M protein (160 kDa neu...	45	0.(
☐	tr	Q97QP7	Immunoglobulin A1 protease [SP1154] [Streptococcus pne...	44	0.(
☐	tr	Q8IB63	Hypothetical protein PF08_0035 [PF08_0035] [Plasmodium...	44	0.(
☐	tr	Q869E1	Similar to Xenopus laevis (African clawed frog). DNA l...	44	0.(
☐	tr	Q9VC00	CG13648-PA [CG13648] [Drosophila melanogaster (Fruit f...	44	0.(
☐	tr	Q839C3	N-acetylmuramoyl-L-alanine amidase, family 4 [EF0252] ...	44	0.(
☐	tr	O73793	Neurofilament medium subunit [Serinus canaria (Canary)]	43	0.(
☐	tr	Q9FN97	Transposon protein-like [Arabidopsis thaliana (Mouse-e...	43	0.(
☐	tr	Q963T1	Glutamate-rich protein (Fragment) [GLURP] [Plasmodium ...	43	0.(
☐	tr	Q07594	K2 protein (Fragment) [Entamoeba histolytica]	43	0.(
☐	tr	Q6PK21	OGFR protein [Homo sapiens (Human)]	43	0.(

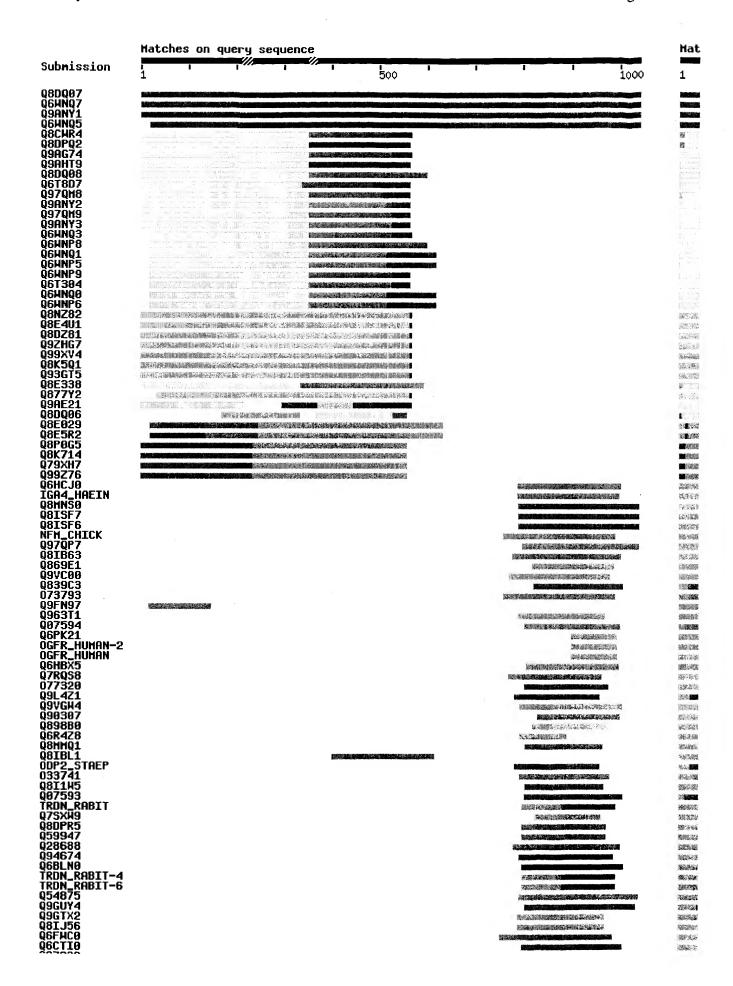|					
□	sp_vs	Q9NZT2-2	Splice isoform 2 of Q9NZT2 [OGFR] [Homo sapiens (...	43	0.0
□	sp	Q9NZT2	OGFR_HUMAN Opioid growth factor receptor (OGFr) (Zeta-...	43	0.0
□	tr	Q6HBX5	Hypothetical protein [BT9727_4640] [Bacillus thuringie...	43	0.0
□	tr	Q7RQS8	Retinitis pigmentosa GTPase regulator-like protein [PY...	43	0.0
□	tr	O77320	Hypothetical protein MAL3P3.3 [MAL3P3.3] [Plasmodium f...	43	0.0
□	tr	Q9L4Z1	Pyruvate dehydrogenase complex subunit E2 [pdhC] [Stap...	42	0.0
□	tr	Q9VGW4	CG14692-PA [CG14692] [Drosophila melanogaster (Fruit f...	42	0.0
□	tr	Q90307	Carassius auratus [Carassius auratus (Goldfish)]	42	0.0
□	tr	Q898B0	Hypothetical protein CTC00555 [CTC00555] [Clostridium ...	42	0.0
□	tr	Q6R4Z8	Dehydrin cor29 [Capsella bursa-pastoris (Shepherd's pu...	42	0.0
□	tr	Q8MMQ1	Similar to Required for the transfer of mannosylphosph...	42	0.0
□	tr	Q8IBL1	Hypothetical protein MAL7P1.129 [MAL7P1.129] [Plasmodi...	42	0.0
□	sp	Q8CT13	ODP2_STAEP Dihydrolipoyllysine-residue acetyltransfera...	41	0.
□	tr	O33741	SpsA protein [Streptococcus pneumoniae]	41	0.
□	tr	Q8I1W5	Hypothetical protein PFD0320c (Fragment) [PFD0320c] [P...	41	0.
□	tr	Q07593	K2 protein (Fragment) [Entamoeba histolytica]	41	0.
□	sp	Q28820	TRDN_RABIT Triadin [TRDN] [Oryctolagus cuniculus (Rabb...	41	0.
□	tr	Q7SXW9	Hypothetical protein (Fragment) [Brachydanio rerio (Ze...	41	0.
□	tr	Q8DPR5	Immunoglobulin A1 protease (EC 3.4.24.13) [iga] [Strep...	41	0.
□	tr	Q59947	Immunoglobulin A1 protease (EC 3.4.24.13) [iga] [Strep...	41	0.
□	tr	Q28688	Neurofilament-H (Fragment) [Oryctolagus cuniculus (Rab...	41	0.
□	tr	Q94674	Thrombospondin-related anonymous protein (Fragment) [T...	41	0.
□	tr	Q6BLN0	Similar to ca|CA2433|IPF12959 Candida albicans IPF1295...	41	0.
□	sp_vs	Q28820-4	Splice isoform Cardiac 3 of Q28820 [TRDN] [Orycto...	41	0.
□	sp_vs	Q28820-6	Splice isoform Skeletal 3 of Q28820 [TRDN] [Oryct...	41	0.
□	tr	Q54875	IgA1 protease [iga] [Streptococcus pneumoniae]	40	0.
□	tr	Q9GUY4	Crustocalcin [DD4(ccn)] [Penaeus japonicus (Kuruma pra...	40	0.
□	tr	Q9GTX2	Glutamate-rich protein [GLURP] [Plasmodium falciparum]	40	0.
□	tr	Q8IJ56	Glutamate-rich protein [PF10_0344] [Plasmodium falcipa...	40	0.
□	tr	Q6FWC0	Candida glabrata strain CBS138 chromosome D complete s...	40	0.
□	tr	Q6CTI0	Similar to sp|Q05050 Saccharomyces cerevisiae YMR031c ...	40	0.
□	tr	Q879S6	Hemagglutinin-like secreted protein [pspA] [Xylella fa...	40	0.
□	tr	Q8IJI0	Pre-mRNA splicing factor, putative [PF10_0217] [Plasmo...	40	0.
□	tr	O96609	Surface antigen ariel1 [Entamoeba histolytica]	40	0.
□	tr	O17102	Hypothetical protein F42G2.6 [F42G2.6] [Caenorhabditis...	40	0.
□	sp_vs	Q28820-5	Splice isoform Skeletal 2 of Q28820 [TRDN] [Oryct...	40	0.
□	sp	P19814	TGN3_RAT Trans-Golgi network integral membrane protein...	40	0.
□	sp	O51498	DPO1_BORBU DNA polymerase I (EC 2.7.7.7) (POL I) [polA...	40	0.
□	sp	P42759	DH10_ARATH Dehydrin ERD10 (Low-temperature-induced pro...	40	0.
□	sp	P13952	CG2B_SPISO G2/mitotic-specific cyclin B [Spisula solid...	40	0.
□	tr	Q63575	Trans golgi network (TGN) specific integral membrane p...	40	0.
□	tr	Q9KK42	Surface protein PspC [pspC] [Streptococcus pneumoniae]	40	0.
□	tr	Q7P6T9	TonB protein [FNV1602] [Fusobacterium nucleatum subsp....	40	0.
□	tr	Q8H7D1	Hypothetical protein [Arabidopsis thaliana (Mouse-ear ...	40	0.
□	tr	Q28687	Neurofilament-H (Fragment) [Oryctolagus cuniculus (Rab...	40	0.

Graphical overview of the alignments

| Click here | to resubmit your query after masking regions matching <u>PROSITE</u> profiles or <u>Pfam</u> HMMs |

(? <u>Help</u>) (use <u>ScanProsite</u> for more details about PROSITE matches)

Profile hits

Pfam hits

Matches on query sequence

Alignments

tr <u>Q8DQ07</u> **Pneumococcal histidine triad protein E [phtE] [Streptococcus** 1039 AA
pneumoniae (strain ATCC BAA-255 / R6)]

<u>align</u>

Score = 2017 bits (5225), Expect = 0.0
Identities = 1004/1039 (96%), Positives = 1004/1039 (96%)


```
Query: 1    MKFSKKYIAAGSAVIVSLSLCAYALNQHRSQENKDNNRVSYVDGSQSSQKSENLTPDQVS 60
            MKFSKKYIAAGSAVIVSLSLCAYALNQHRSQENKDNNRVSYVDGSQSSQKSENLTPDQVS
Sbjct: 1    MKFSKKYIAAGSAVIVSLSLCAYALNQHRSQENKDNNRVSYVDGSQSSQKSENLTPDQVS 60

Query: 61   QKEGIQAEQIVIKITDQGYVTSHGDHYHYYNGKVPYDALFSEELLMKDPNYQLKDADIVN 120
            QKEGIQAEQIVIKITDQGYVTSHGDHYHYYNGKVPYDALFSEELLMKDPNYQLKDADIVN
Sbjct: 61   QKEGIQAEQIVIKITDQGYVTSHGDHYHYYNGKVPYDALFSEELLMKDPNYQLKDADIVN 120

Query: 121  EVKGGYIIKVDGKYYVYLKDAAHADNVRTKDEINRQKQEHVKDNEKVNSNVAVARSQGRY 180
            EVKGGYIIKVDGKYYVYLKDAAHADNVRTKDEINRQKQEHVKDNEKVNSNVAVARSQGRY
Sbjct: 121  EVKGGYIIKVDGKYYVYLKDAAHADNVRTKDEINRQKQEHVKDNEKVNSNVAVARSQGRY 180

Query: 181  TTNDGYVFNPADIIEDTGNAYIVPHGGHYHYIPXXXXXXXXXXXXXXXXXXXXXNMQPSQLS 240
            TTNDGYVFNPADIIEDTGNAYIVPHGGHYHYIP                     NMQPSQLS
Sbjct: 181  TTNDGYVFNPADIIEDTGNAYIVPHGGHYHYIPKSDLSASELAAAKAHLAGKNMQPSQLS 240

Query: 241  YSSTASDNNTQSVAKGSTSKPANKSENLQSLLKELYDSPSAQRYSESDGLVFDPAKIISR 300
            YSSTASDNNTQSVAKGSTSKPANKSENLQSLLKELYDSPSAQRYSESDGLVFDPAKIISR
Sbjct: 241  YSSTASDNNTQSVAKGSTSKPANKSENLQSLLKELYDSPSAQRYSESDGLVFDPAKIISR 300

Query: 301  TPNGVAIPHGDHYHFIPYSKLSALEEKIARRVPISGTGSTVSTNAKPNEVVXXXXXXXXX 360
            TPNGVAIPHGDHYHFIPYSKLSALEEKIARRVPISGTGSTVSTNAKPNEVV
Sbjct: 301  TPNGVAIPHGDHYHFIPYSKLSALEEKIARRVPISGTGSTVSTNAKPNEVVSSLGSLSSN 360

Query: 361  XXXXXXXXKELSSASDGYIFNPKDIVEETATAYIVRHGDHFHYIPKSNQIGQPTLPNNSLA 420
                    KELSSASDGYIFNPKDIVEETATAYIVRHGDHFHYIPKSNQIGQPTLPNNSLA
Sbjct: 361  PSSLTTSKELSSASDGYIFNPKDIVEETATAYIVRHGDHFHYIPKSNQIGQPTLPNNSLA 420

Query: 421  TPSPSLPINPGISHEKHEEDGYGFDANRIIAEDESGFIMSHGNHNHYFFKKDLTEEQIKA 480
            TPSPSLPINPGISHEKHEEDGYGFDANRIIAEDESGFIMSHGNHNHYFFKKDLTEEQIKA
Sbjct: 421  TPSPSLPINPGISHEKHEEDGYGFDANRIIAEDESGFIMSHGNHNHYFFKKDLTEEQIKA 480

Query: 481  AQKHLEEVKTSHNGLDSLSSHEQDYPGNAKEMKDLDKKIEEKIAGIMKQYGVKRESIVVN 540
            AQKHLEEVKTSHNGLDSLSSHEQDYPGNAKEMKDLDKKIEEKIAGIMKQYGVKRESIVVN
Sbjct: 481  AQKHLEEVKTSHNGLDSLSSHEQDYPGNAKEMKDLDKKIEEKIAGIMKQYGVKRESIVVN 540

Query: 541  KEKNAIIYPHGDHHHADPIDEHKPVGIGHSHSNYELFKPEEGVAKKEGNKVYTGEELTNV 600
            KEKNAIIYPHGDHHHADPIDEHKPVGIGHSHSNYELFKPEEGVAKKEGNKVYTGEELTNV
Sbjct: 541  KEKNAIIYPHGDHHHADPIDEHKPVGIGHSHSNYELFKPEEGVAKKEGNKVYTGEELTNV 600

Query: 601  VNLLKNSTFNNQNFTLANGQKRVSFSFPPELEKKLGINMLVKLITPDGKVLEKVSGKVFG 660
            VNLLKNSTFNNQNFTLANGQKRVSFSFPPELEKKLGINMLVKLITPDGKVLEKVSGKVFG
Sbjct: 601  VNLLKNSTFNNQNFTLANGQKRVSFSFPPELEKKLGINMLVKLITPDGKVLEKVSGKVFG 660

Query: 661  EGVGNIANFELDQPYLPGQTFKYTIASKDYPEVSYDGTFTVPTSLAYKMASQTIFYPFHA 720
            EGVGNIANFELDQPYLPGQTFKYTIASKDYPEVSYDGTFTVPTSLAYKMASQTIFYPFHA
Sbjct: 661  EGVGNIANFELDQPYLPGQTFKYTIASKDYPEVSYDGTFTVPTSLAYKMASQTIFYPFHA 720
```

```
Query:  721  GDTYLRVNPQFAVPKGTDALVRVFDEFHGNAYLENNYKVGEIKLPIPKLNQGTTRTAGNK 780
             GDTYLRVNPQFAVPKGTDALVRVFDEFHGNAYLENNYKVGEIKLPIPKLNQGTTRTAGNK
Sbjct:  721  GDTYLRVNPQFAVPKGTDALVRVFDEFHGNAYLENNYKVGEIKLPIPKLNQGTTRTAGNK 780

Query:  781  IPVTFMANAYLDNQSTYIVEVPILEKENQTDKPSILPQFKRNKAQENSKLDEKVEEPKTS 840
             IPVTFMANAYLDNQSTYIVEVPILEKENQTDKPSILPQFKRNKAQENSKLDEKVEEPKTS
Sbjct:  781  IPVTFMANAYLDNQSTYIVEVPILEKENQTDKPSILPQFKRNKAQENSKLDEKVEEPKTS 840

Query:  841  EKVEKEKLSETGNSTSNSTLEEVPTVDPVQEKVAKFAESYGMKLENVLFNMDGTIELYLP 900
             EKVEKEKLSETGNSTSNSTLEEVPTVDPVQEKVAKFAESYGMKLENVLFNMDGTIELYLP
Sbjct:  841  EKVEKEKLSETGNSTSNSTLEEVPTVDPVQEKVAKFAESYGMKLENVLFNMDGTIELYLP 900

Query:  901  SGEVIKKNMADFTGEAPQGNGENKPSENGKVSTGTVENQPTENKPADSLPEAPNEKPVKP 960
             SGEVIKKNMADFTGEAPQGNGENKPSENGKVSTGTVENQPTENKPADSLPEAPNEKPVKP
Sbjct:  901  SGEVIKKNMADFTGEAPQGNGENKPSENGKVSTGTVENQPTENKPADSLPEAPNEKPVKP 960

Query:  961  ENSTDNGMLNPEGNVGSDPMLDPALEEAPAVDPVQEKLEKFTASYGLGLDSVIFNMDGTI 1020
             ENSTDNGMLNPEGNVGSDPMLDPALEEAPAVDPVQEKLEKFTASYGLGLDSVIFNMDGTI
Sbjct:  961  ENSTDNGMLNPEGNVGSDPMLDPALEEAPAVDPVQEKLEKFTASYGLGLDSVIFNMDGTI 1020

Query: 1021  ELRLPSGEVIKKNLSDLIA 1039
             ELRLPSGEVIKKNLSDLIA
Sbjct: 1021  ELRLPSGEVIKKNLSDLIA 1039
```

tr   Q6WNQ7   **Surface protein BVH-3 [bvh-3] [Streptococcus   1039 AA
                 pneumoniae]**   align

```
 Score = 2017 bits (5225), Expect = 0.0
 Identities = 1004/1039 (96%), Positives = 1004/1039 (96%)

Query:   1   MKFSKKYIAAGSAVIVSLSLCAYALNQHRSQENKDNNRVSYVDGSQSSQKSENLTPDQVS 60
             MKFSKKYIAAGSAVIVSLSLCAYALNQHRSQENKDNNRVSYVDGSQSSQKSENLTPDQVS
Sbjct:   1   MKFSKKYIAAGSAVIVSLSLCAYALNQHRSQENKDNNRVSYVDGSQSSQKSENLTPDQVS 60

Query:  61   QKEGIQAEQIVIKITDQGYVTSHGDHYHYYNGKVPYDALFSEELLMKDPNYQLKDADIVN 120
             QKEGIQAEQIVIKITDQGYVTSHGDHYHYYNGKVPYDALFSEELLMKDPNYQLKDADIVN
Sbjct:  61   QKEGIQAEQIVIKITDQGYVTSHGDHYHYYNGKVPYDALFSEELLMKDPNYQLKDADIVN 120

Query: 121   EVKGGYIIKVDGKYYVYLKDAAHADNVRTKDEINRQKQEHVKDNEKVNSNVAVARSQGRY 180
             EVKGGYIIKVDGKYYVYLKDAAHADNVRTKDEINRQKQEHVKDNEKVNSNVAVARSQGRY
Sbjct: 121   EVKGGYIIKVDGKYYVYLKDAAHADNVRTKDEINRQKQEHVKDNEKVNSNVAVARSQGRY 180

Query: 181   TTNDGYVFNPADIIEDTGNAYIVPHGGHYHYIPXXXXXXXXXXXXXXXXXXXXXNMQPSQLS 240
             TTNDGYVFNPADIIEDTGNAYIVPHGGHYHYIP                     NMQPSQLS
Sbjct: 181   TTNDGYVFNPADIIEDTGNAYIVPHGGHYHYIPKSDLSASELAAAKAHLAGKNMQPSQLS 240

Query: 241   YSSTASDNNTQSVAKGSTSKPANKSENLQSLLKELYDSPSAQRYSESDGLVFDPAKIISR 300
             YSSTASDNNTQSVAKGSTSKPANKSENLQSLLKELYDSPSAQRYSESDGLVFDPAKIISR
Sbjct: 241   YSSTASDNNTQSVAKGSTSKPANKSENLQSLLKELYDSPSAQRYSESDGLVFDPAKIISR 300

Query: 301   TPNGVAIPHGDHYHFIPYSKLSALEEKIARRVPISGTGSTVSTNAKPNEVVXXXXXXXXX 360
             TPNGVAIPHGDHYHFIPYSKLSALEEKIARRVPISGTGSTVSTNAKPNEVV
Sbjct: 301   TPNGVAIPHGDHYHFIPYSKLSALEEKIARRVPISGTGSTVSTNAKPNEVVSSLGSLSSN 360

Query: 361   XXXXXXXKELSSASDGYIFNPKDIVEETATAYIVRHGDHFHYIPKSNQIGQPTLPNNSLA 420
```

```
                 KELSSASDGYIFNPKDIVEETATAYIVRHGDHFHYIPKSNQIGQPTLPNNSLA
Sbjct: 361  PSSLTTSKELSSASDGYIFNPKDIVEETATAYIVRHGDHFHYIPKSNQIGQPTLPNNSLA 420

Query: 421  TPSPSLPINPGISHEKHEEDGYGFDANRIIAEDESGFIMSHGNHNHYFFKKDLTEEQIKA 480
            TPSPSLPINPGISHEKHEEDGYGFDANRIIAEDESGFIMSHGNHNHYFFKKDLTEEQIKA
Sbjct: 421  TPSPSLPINPGISHEKHEEDGYGFDANRIIAEDESGFIMSHGNHNHYFFKKDLTEEQIKA 480

Query: 481  AQKHLEEVKTSHNGLDSLSSHEQDYPGNAKEMKDLDKKIEEKIAGIMKQYGVKRESIVVN 540
            AQKHLEEVKTSHNGLDSLSSHEQDYPGNAKEMKDLDKKIEEKIAGIMKQYGVKRESIVVN
Sbjct: 481  AQKHLEEVKTSHNGLDSLSSHEQDYPGNAKEMKDLDKKIEEKIAGIMKQYGVKRESIVVN 540

Query: 541  KEKNAIIYPHGDHHHADPIDEHKPVGIGHSHSNYELFKPEEGVAKKEGNKVYTGEELTNV 600
            KEKNAIIYPHGDHHHADPIDEHKPVGIGHSHSNYELFKPEEGVAKKEGNKVYTGEELTNV
Sbjct: 541  KEKNAIIYPHGDHHHADPIDEHKPVGIGHSHSNYELFKPEEGVAKKEGNKVYTGEELTNV 600

Query: 601  VNLLKNSTFNNQNFTLANGQKRVSFSFPPELEKKLGINMLVKLITPDGKVLEKVSGKVFG 660
            VNLLKNSTFNNQNFTLANGQKRVSFSFPPELEKKLGINMLVKLITPDGKVLEKVSGKVFG
Sbjct: 601  VNLLKNSTFNNQNFTLANGQKRVSFSFPPELEKKLGINMLVKLITPDGKVLEKVSGKVFG 660

Query: 661  EGVGNIANFELDQPYLPGQTFKYTIASKDYPEVSYDGTFTVPTSLAYKMASQTIFYPFHA 720
            EGVGNIANFELDQPYLPGQTFKYTIASKDYPEVSYDGTFTVPTSLAYKMASQTIFYPFHA
Sbjct: 661  EGVGNIANFELDQPYLPGQTFKYTIASKDYPEVSYDGTFTVPTSLAYKMASQTIFYPFHA 720

Query: 721  GDTYLRVNPQFAVPKGTDALVRVFDEFHGNAYLENNYKVGEIKLPIPKLNQGTTRTAGNK 780
            GDTYLRVNPQFAVPKGTDALVRVFDEFHGNAYLENNYKVGEIKLPIPKLNQGTTRTAGNK
Sbjct: 721  GDTYLRVNPQFAVPKGTDALVRVFDEFHGNAYLENNYKVGEIKLPIPKLNQGTTRTAGNK 780

Query: 781  IPVTFMANAYLDNQSTYIVEVPILEKENQTDKPSILPQFKRNKAQENSKLDEKVEEPKTS 840
            IPVTFMANAYLDNQSTYIVEVPILEKENQTDKPSILPQFKRNKAQENSKLDEKVEEPKTS
Sbjct: 781  IPVTFMANAYLDNQSTYIVEVPILEKENQTDKPSILPQFKRNKAQENSKLDEKVEEPKTS 840

Query: 841  EKVEKEKLSETGNSTSNSTLEEVPTVDPVQEKVAKFAESYGMKLENVLFNMDGTIELYLP 900
            EKVEKEKLSETGNSTSNSTLEEVPTVDPVQEKVAKFAESYGMKLENVLFNMDGTIELYLP
Sbjct: 841  EKVEKEKLSETGNSTSNSTLEEVPTVDPVQEKVAKFAESYGMKLENVLFNMDGTIELYLP 900

Query: 901  SGEVIKKNMADFTGEAPQGNGENKPSENGKVSTGTVENQPTENKPADSLPEAPNEKPVKP 960
            SGEVIKKNMADFTGEAPQGNGENKPSENGKVSTGTVENQPTENKPADSLPEAPNEKPVKP
Sbjct: 901  SGEVIKKNMADFTGEAPQGNGENKPSENGKVSTGTVENQPTENKPADSLPEAPNEKPVKP 960

Query: 961  ENSTDNGMLNPEGNVGSDPMLDPALEEAPAVDPVQEKLEKFTASYGLGLDSVIFNMDGTI 1020
            ENSTDNGMLNPEGNVGSDPMLDPALEEAPAVDPVQEKLEKFTASYGLGLDSVIFNMDGTI
Sbjct: 961  ENSTDNGMLNPEGNVGSDPMLDPALEEAPAVDPVQEKLEKFTASYGLGLDSVIFNMDGTI 1020

Query: 1021 ELRLPSGEVIKKNLSDLIA 1039
            ELRLPSGEVIKKNLSDLIA
Sbjct: 1021 ELRLPSGEVIKKNLSDLIA 1039
```

tr Q9ANY1 **Pneumococcal histidine triad protein E precursor**
          (Hypothetical
          protein SP1004) [phtE] [Streptococcus pneumoniae]

1039
AA
align

```
 Score = 2006 bits (5196), Expect = 0.0
 Identities = 998/1039 (96%), Positives = 1000/1039 (96%)

Query: 1    MKFSKKYIAAGSAVIVSLSLCAYALNQHRSQENKDNNRVSYVDGSQSSQKSENLTPDQVS 60
```

```
                MKFSKKYIAAGSAVIVSLSLCAYALNQHRSQENKDNNRVSYVDGSQSSQKSENLTPDQVS
Sbjct: 1        MKFSKKYIAAGSAVIVSLSLCAYALNQHRSQENKDNNRVSYVDGSQSSQKSENLTPDQVS 60

Query: 61       QKEGIQAEQIVIKITDQGYVTSHGDHYHYYNGKVPYDALFSEELLMKDPNYQLKDADIVN 120
                QKEGIQAEQIVIKITDQGYVTSHGDHYHYYNGKVPYDALFSEELLMKDPNYQLKDADIVN
Sbjct: 61       QKEGIQAEQIVIKITDQGYVTSHGDHYHYYNGKVPYDALFSEELLMKDPNYQLKDADIVN 120

Query: 121      EVKGGYIIKVDGKYYVYLKDAAHADNVRTKDEINRQKQEHVKDNEKVNSNVAVARSQGRY 180
                EVKGGYIIKVDGKYYVYLKDAAHADNVRTKDEINRQKQEHVKDNEKVNSNVAVARSQGRY
Sbjct: 121      EVKGGYIIKVDGKYYVYLKDAAHADNVRTKDEINRQKQEHVKDNEKVNSNVAVARSQGRY 180

Query: 181      TTNDGYVFNPADIIEDTGNAYIVPHGGHYHYIPXXXXXXXXXXXXXXXXXXXXXXNMQPSQLS 240
                TTNDGYVFNPADIIEDTGNAYIVPHGGHYHYIP                     NMQPSQLS
Sbjct: 181      TTNDGYVFNPADIIEDTGNAYIVPHGGHYHYIPKSDLSASELAAAKAHLAGKNMQPSQLS 240

Query: 241      YSSTASDNNTQSVAKGSTSKPANKSENLQSLLKELYDSPSAQRYSESDGLVFDPAKIISR 300
                YSSTASDNNTQSVAKGSTSKPANKSENLQSLLKELYDSPSAQRYSESDGLVFDPAKIISR
Sbjct: 241      YSSTASDNNTQSVAKGSTSKPANKSENLQSLLKELYDSPSAQRYSESDGLVFDPAKIISR 300

Query: 301      TPNGVAIPHGDHYHFIPYSKLSALEEKIARRVPISGTGSTVSTNAKPNEVVXXXXXXXXX 360
                TPNGVAIPHGDHYHFIPYSKLSALEEKIAR VPISGTGSTVSTNAKPNEVV
Sbjct: 301      TPNGVAIPHGDHYHFIPYSKLSALEEKIARMVPISGTGSTVSTNAKPNEVVSSLGSLSSN 360

Query: 361      XXXXXXXKELSSASDGYIFNPKDIVEETATAYIVRHGDHFHYIPKSNQIGQPTLPNNSLA 420
                       KELSSASDGYIFNPKDIVEETATAYIVRHGDHFHYIPKSNQIGQPTLPNNSLA
Sbjct: 361      PSSLTTSKELSSASDGYIFNPKDIVEETATAYIVRHGDHFHYIPKSNQIGQPTLPNNSLA 420

Query: 421      TPSPSLPINPGISHEKHEEDGYGFDANRIIAEDESGFIMSHGNHNHYFFKKDLTEEQIKA 480
                TPSPSLPINPG SHEKHEEDGYGFDANRIIAEDESGF+MSHG+HNHYFFKKDLTEEQIKA
Sbjct: 421      TPSPSLPINPGTSHEKHEEDGYGFDANRIIAEDESGFVMSHGDHNHYFFKKDLTEEQIKA 480

Query: 481      AQKHLEEVKTSHNGLDSLSSHEQDYPGNAKEMKDLDKKIEEKIAGIMKQYGVKRESIVVN 540
                AQKHLEEVKTSHNGLDSLSSHEQDYP NAKEMKDLDKKIEEKIAGIMKQYGVKRESIVVN
Sbjct: 481      AQKHLEEVKTSHNGLDSLSSHEQDYPSNAKEMKDLDKKIEEKIAGIMKQYGVKRESIVVN 540

Query: 541      KEKNAIIYPHGDHHHADPIDEHKPVGIGHSHSNYELFKPEEGVAKKEGNKVYTGEELTNV 600
                KEKNAIIYPHGDHHHADPIDEHKPVGIGHSHSNYELFKPEEGVAKKEGNKVYTGEELTNV
Sbjct: 541      KEKNAIIYPHGDHHHADPIDEHKPVGIGHSHSNYELFKPEEGVAKKEGNKVYTGEELTNV 600

Query: 601      VNLLKNSTFNNQNFTLANGQKRVSFSFPPELEKKLGINMLVKLITPDGKVLEKVSGKVFG 660
                VNLLKNSTFNNQNFTLANGQKRVSFSFPPELEKKLGINMLVKLITPDGKVLEKVSGKVFG
Sbjct: 601      VNLLKNSTFNNQNFTLANGQKRVSFSFPPELEKKLGINMLVKLITPDGKVLEKVSGKVFG 660

Query: 661      EGVGNIANFELDQPYLPGQTFKYTIASKDYPEVSYDGTFTVPTSLAYKMASQTIFYPFHA 720
                EGVGNIANFELDQPYLPGQTFKYTIASKDYPEVSYDGTFTVPTSLAYKMASQTIFYPFHA
Sbjct: 661      EGVGNIANFELDQPYLPGQTFKYTIASKDYPEVSYDGTFTVPTSLAYKMASQTIFYPFHA 720

Query: 721      GDTYLRVNPQFAVPKGTDALVRVFDEFHGNAYLENNYKVGEIKLPIPKLNQGTTRTAGNK 780
                GDTYLRVNPQFAVPKGTDALVRVFDEFHGNAYLENNYKVGEIKLPIPKLNQGTTRTAGNK
Sbjct: 721      GDTYLRVNPQFAVPKGTDALVRVFDEFHGNAYLENNYKVGEIKLPIPKLNQGTTRTAGNK 780

Query: 781      IPVTFMANAYLDNQSTYIVEVPILEKENQTDKPSILPQFKRNKAQENSKLDEKVEEPKTS 840
                IPVTFMANAYLDNQSTYIVEVPILEKENQTDKPSILPQFKRNKAQEN KLDEKVEEPKTS
Sbjct: 781      IPVTFMANAYLDNQSTYIVEVPILEKENQTDKPSILPQFKRNKAQENLKLDEKVEEPKTS 840

Query: 841      EKVEKEKLSETGNSTSNSTLEEVPTVDPVQEKVAKFAESYGMKLENVLFNMDGTIELYLP 900
                EKVEKEKLSETGNSTSNSTLEEVPTVDPVQEKVAKFAESYGMKLENVLFNMDGTIELYLP
Sbjct: 841      EKVEKEKLSETGNSTSNSTLEEVPTVDPVQEKVAKFAESYGMKLENVLFNMDGTIELYLP 900
```

```
Query:  901  SGEVIKKNMADFTGEAPQGNGENKPSENGKVSTGTVENQPTENKPADSLPEAPNEKPVKP 960
             SGEVIKKNMADFTGEAPQGNGENKPSENGKVSTGTVENQPTENKPADSLPEAPNEKPVKP
Sbjct:  901  SGEVIKKNMADFTGEAPQGNGENKPSENGKVSTGTVENQPTENKPADSLPEAPNEKPVKP 960

Query:  961  ENSTDNGMLNPEGNVGSDPMLDPALEEAPAVDPVQEKLEKFTASYGLGLDSVIFNMDGTI 1020
             ENSTDNGMLNPEGNVGSDPMLDPALEEAPAVDPVQEKLEKFTASYGLGLDSVIFNMDGTI
Sbjct:  961  ENSTDNGMLNPEGNVGSDPMLDPALEEAPAVDPVQEKLEKFTASYGLGLDSVIFNMDGTI 1020

Query: 1021  ELRLPSGEVIKKNLSDLIA 1039
             ELRLPSGEVIKKNLSDLIA
Sbjct: 1021  ELRLPSGEVIKKNLSDLIA 1039
```

tr Q6WNQ5 **Surface protein BVH-3 (Fragment) [bvh-3] [Streptococcus pneumoniae]**   1019
                                                                                      AA
                                                                                      align

```
 Score = 1968 bits (5099), Expect = 0.0
 Identities = 977/1019 (95%), Positives = 979/1019 (95%)

Query:   21  CAYALNQHRSQENKDNNRVSYVDGSQSSQKSENLTPDQVSQKEGIQAEQIVIKITDQGYV 80
             CAYALNQHRSQENKDNNRVSYVDGSQSSQKSENLTPDQVSQKEGIQAEQIVIKITDQGYV
Sbjct:    1  CAYALNQHRSQENKDNNRVSYVDGSQSSQKSENLTPDQVSQKEGIQAEQIVIKITDQGYV 60

Query:   81  TSHGDHYHYYNGKVPYDALFSEELLMKDPNYQLKDADIVNEVKGGYIIKVDGKYYVYLKD 140
             TSHGDHYHYYNGKVPYDALFSEELLMKDPNYQLKDADIVNEVKGGYIIKVDGKYYVYLKD
Sbjct:   61  TSHGDHYHYYNGKVPYDALFSEELLMKDPNYQLKDADIVNEVKGGYIIKVDGKYYVYLKD 120

Query:  141  AAHADNVRTKDEINRQKQEHVKDNEKVNSNVAVARSQGRYTTNDGYVFNPADIIEDTGNA 200
             AAHADNVRTKDEINRQKQEHVKDNEKVNSNVAVARSQGRYTTNDGYVFNPADIIEDTGNA
Sbjct:  121  AAHADNVRTKDEINRQKQEHVKDNEKVNSNVAVARSQGRYTTNDGYVFNPADIIEDTGNA 180

Query:  201  YIVPHGGHYHYIPXXXXXXXXXXXXXXXXXXXXXXXXNMQPSQLSYSSTASDNNTQSVAKGSTSK 260
             YIVPH GHYHYIP                        NMQPSQLSYSSTASDNNTQSVAKGSTSK
Sbjct:  181  YIVPHRGHYHYIPKSDLSASELAAAKAHLAGKNMQPSQLSYSSTASDNNTQSVAKGSTSK 240

Query:  261  PANKSENLQSLLKELYDSPSAQRYSESDGLVFDPAKIISRTPNGVAIPHGDHYHFIPYSK 320
             PANKSENLQSLLKELYDSPSAQRYSESDGLVFDPAKIISRTPNGVAIPHGDHYHFIPYSK
Sbjct:  241  PANKSENLQSLLKELYDSPSAQRYSESDGLVFDPAKIISRTPNGVAIPHGDHYHFIPYSK 300

Query:  321  LSALEEKIARRVPISGTGSTVSTNAKPNEVVXXXXXXXXXXXXXXXXXXXXKELSSASDGYIFN 380
             LSALEEKIAR VPISGTGSTVSTNAKPNEVV                    KELSSASDGYIFN
Sbjct:  301  LSALEEKIARMVPISGTGSTVSTNAKPNEVVSSLGSLSSNPSSLTTSKELSSASDGYIFN 360

Query:  381  PKDIVEETATAYIVRHGDHFHYIPKSNQIGQPTLPNNSLATPSPSLPINPGISHEKHEED 440
             PKDIVEETATAYIVRHGDHFHYIPKSNQIGQPTLPNNSLATPSPSLPINPG SHEKHEED
Sbjct:  361  PKDIVEETATAYIVRHGDHFHYIPKSNQIGQPTLPNNSLATPSPSLPINPGTSHEKHEED 420

Query:  441  GYGFDANRIIAEDESGFIMSHGNHNHYFFKKDLTEEQIKAAQKHLEEVKTSHNGLDSLSS 500
             GYGFDANRIIAEDESGF+MSHG+HNHYFFKKDLTEEQIKAAQKHLEEVKTSHNGLDSLSS
Sbjct:  421  GYGFDANRIIAEDESGFVMSHGDHNHYFFKKDLTEEQIKAAQKHLEEVKTSHNGLDSLSS 480

Query:  501  HEQDYPGNAKEMKDLDKKIEEKIAGIMKQYGVKRESIVVNKEKNAIIYPHGDHHHADPID 560
             HEQDYP NAKEMKDLDKKIEEKIAGIMKQYGVKRESIVVNKEKNAIIYPHGDHHHADPID
Sbjct:  481  HEQDYPSNAKEMKDLDKKIEEKIAGIMKQYGVKRESIVVNKEKNAIIYPHGDHHHADPID 540
```

```
Query:  561 EHKPVGIGHSHSNYELFKPEEGVAKKEGNKVYTGEELTNVVNLLKNSTFNNQNFTLANGQ 620
            EHKPVGIGHSHSNYELFKPEEGVAKKEGNKVYTGEELTNVVNLLKNSTFNNQNFTLANGQ
Sbjct:  541 EHKPVGIGHSHSNYELFKPEEGVAKKEGNKVYTGEELTNVVNLLKNSTFNNQNFTLANGQ 600

Query:  621 KRVSFSFPPELEKKLGINMLVKLITPDGKVLEKVSGKVFGEGVGNIANFELDQPYLPGQT 680
            KRVSFSFPPELEKKLGINMLVKLITPDGKVLEKVSGKVFGEGVGNIANFELDQPYLPGQT
Sbjct:  601 KRVSFSFPPELEKKLGINMLVKLITPDGKVLEKVSGKVFGEGVGNIANFELDQPYLPGQT 660

Query:  681 FKYTIASKDYPEVSYDGTFTVPTSLAYKMASQTIFYPFHAGDTYLRVNPQFAVPKGTDAL 740
            FKYTIASKDYPEVSYDGTFTVPTSLAYKMASQTIFYPFHAGDTYLRVNPQFAVPKGTDAL
Sbjct:  661 FKYTIASKDYPEVSYDGTFTVPTSLAYKMASQTIFYPFHAGDTYLRVNPQFAVPKGTDAL 720

Query:  741 VRVFDEFHGNAYLENNYKVGEIKLPIPKLNQGTTRTAGNKIPVTFMANAYLDNQSTYIVE 800
            VRVFDEFHGNAYLENNYKVGEIKLPIPKLNQGTTRTAGNKIPVTFMANAYLDNQSTYIVE
Sbjct:  721 VRVFDEFHGNAYLENNYKVGEIKLPIPKLNQGTTRTAGNKIPVTFMANAYLDNQSTYIVE 780

Query:  801 VPILEKENQTDKPSILPQFKRNKAQENSKLDEKVEEPKTSEKVEKEKLSETGNSTSNSTL 860
            VPILEKENQTDKPSILPQFKRNKAQENSK DEKVEEPKTSEKVEKEKLSETGNSTSNSTL
Sbjct:  781 VPILEKENQTDKPSILPQFKRNKAQENSKFDEKVEEPKTSEKVEKEKLSETGNSTSNSTL 840

Query:  861 EEVPTVDPVQEKVAKFAESYGMKLENVLFNMDGTIELYLPSGEVIKKNMADFTGEAPQGN 920
            EEVPTVDPVQEKVAKFAESYGMKLENVLFNMDGTIELYLPSGEVIKKNMADFTGEAPQGN
Sbjct:  841 EEVPTVDPVQEKVAKFAESYGMKLENVLFNMDGTIELYLPSGEVIKKNMADFTGEAPQGN 900

Query:  921 GENKPSENGKVSTGTVENQPTENKPADSLPEAPNEKPVKPENSTDNGMLNPEGNVGSDPM 980
            GENKPSENGKVSTGTVENQPTENKPADSLPEAPNEKPVKPENSTDNGMLNPEGNVGSDPM
Sbjct:  901 GENKPSENGKVSTGTVENQPTENKPADSLPEAPNEKPVKPENSTDNGMLNPEGNVGSDPM 960

Query:  981 LDPALEEAPAVDPVQEKLEKFTASYGLGLDSVIFNMDGTIELRLPSGEVIKKNLSDLIA 1039
            LDPALEEAPAVDPVQEKLEKFTASYGLGLDSVIFNMDGTIELRLPSGEVIKKNLSDLIA
Sbjct:  961 LDPALEEAPAVDPVQEKLEKFTASYGLGLDSVIFNMDGTIELRLPSGEVIKKNLSDLIA 1019
```

tr <u>Q8CWR4</u> **Histidine Motif-Containing protein [phpA] [Streptococcus**   855
       pneumoniae   AA
       (strain ATCC BAA-255 / R6)]   <u>align</u>

```
 Score =  442 bits (1137), Expect = e-122
 Identities = 219/369 (59%), Positives = 271/369 (73%), Gaps = 21/369 (5%)

Query:    1 MKFSKKYIAAGSAVIVSLSLCAYALNQHRSQENKDNNRVSYVDGSQSSQKSENLTPDQVS 60
            MK +KKY+A   A +V LS+CAY L  H++Q  K+NNRVSY+DG Q++QK+ENLTPD+VS
Sbjct:   12 MKINKKYLAGSVATLV-LSVCAYELGLHQAQTVKENNRVSYIDGKQATQKTENLTPDEVS 70

Query:   61 QKEGIQAEQIVIKITDQGYVTSHGDHYHYYNGKVPYDALFSEELLMKDPNYQLKDADIVN 120
            ++EGI AEQIVIKITDQGYVTSHGDHYHYYNGKVPYDA+ SEELLMKDPNYQLKD DI++
Sbjct:   71 KREGINAEQIVIKITDQGYVTSHGDHYHYYNGKVPYDAIISEELLMKDPNYQLKDEDIIS 130

Query:  121 EVKGGYIIKVDGKYYVYLKDAAHADNVRTKDEINRQKQEHVKDNE----KVNSNVAVARS 176
            E+KGGY+IKVDGKYYVYLKDAAHADNVRTK+EINRQKQEH +  E    + +  VA+ARS
Sbjct:  131 EIKGGYVIKVDGKYYVYLKDAAHADNVRTKEEINRQKQEHSQHREGGTPRNDGAVALARS 190

Query:  177 QGRYTTNDGYVFNPADIIEDTGNAYIVPHGGHYHYIPXXXXXXXXXXXXXXXXXXXXXNMQP 236
            QGRYTT+DGY+FN +DIIEDTG+AYIVPHG HYHYIP
Sbjct:  191 QGRYTTDDGYIFNASDIIEDTGDAYIVPHGDHYHYIPKNELSASELAAAKAFLSGRGNLS 250
```

```
Query:  237  SQLSYSSTASDN---------------NTQSVAKGSTSKPANKSENLQSLLKELYDSPS  280
              +  +Y   SDN             NT +    +T+  A++S ++  SLLK+LY  P
Sbjct:  251  NSRTYRRQNSDNTSRTNWVPSVSNPGTTNTNTSNNSNTNSQASQSNDIDSLLKQLYKLPL  310

Query:  281  AQRYSESDGLVFDPAKIISRTPNGVAIPHGDHYHFIPYSKLSALEEKIARRVPISGTGST  340
              +QR+ ESDGL+FDPA+I SRT NGVA+PHGDHYHFIPYS+LS LEEK+AR +P+     +
Sbjct:  311  SQRHVESDGLIFDPAQITSRTANGVAVPHGDHYHFIPYSQLSPLEEKLARIIPLRYRSNH  370

Query:  341  VSTNAKPNE  349
              +++P +
Sbjct:  371  WVPDSRPEQ  379


 Score = 84.0 bits (206), Expect = 1e-14
 Identities = 111/511 (21%), Positives = 190/511 (36%), Gaps = 70/511 (13%)

Query:  71   VIKITDQGYVTSHGDHYHYYNGKVPYDALFSEEL-----------LMKDPNYQLKDADI  118
              +I+ T  Y+  HGDHYHY    +P + L + EL          L    Y+ +++D
Sbjct:  207  IIEDTGDAYIVPHGDHYHY----IPKNELSASELAAAKAFLSGRGNLSNSRTYRRQNSD-  261

Query:  119  VNEVKGGYIIKVDGKYYVYLKDAAHADNVRTKDEINRQKQ-EHVKDNEKVNSNVAVARSQ  177
               N +  ++ V         N  T +  N  Q    D + +
Sbjct:  262  -NTSRTNWVPSVSNP---------GTTNTNTSNNSNTNSQASQSNDIDSLLKQLYKLPLS  311

Query:  178  GRYTTNDGYVFNPADIIEDTGNAYIVPHGGHYHYIPXXXXXXXXXXXXXXXXXXXX-----  232
              R+  +DG +F+PA I   T N   VPHG HYH+IP
Sbjct:  312  QRHVESDGLIFDPAQITSRTANGVAVPHGDHYHFIPYSQLSPLEEKLARIIPLRYRSNHW  371

Query:  233  --NMQPSQLSYSSTASDNNTQSVAKGSTSKPANKSENLQSLLKELYDSPSAQRYSESDGL  290
              + +P Q S  ST  + +   A      P+N  +  + L+KE            E +G+
Sbjct:  372  VPDSRPEQPSPQSTPEPSPSPQPAPNPQPAPSNPID--EKLVKEAVRKVGDGYVFEENGV  429

Query:  291  V-FDPAKIIS-RTPNGVAIPHGDHYHFIPYSKLSALEEKIARRVPISGTGSTVSTNAKPN  348
              + PAK +S  T  G+           SKL A +E ++ ++    T    S    N
Sbjct:  430  PRYIPAKDLSAETAAGID-----------SKL-AKQESLSHKLGAKKTDLPSSDREFYN  476

Query:  349  EVVXXXXXXXXXXXXXXXXXXXKELSSASDGYIFNPKDIVEETATAYIVRHGDHFHYIPKSNQ  408
              +                  +   A D +  KD+ +         D  ++
Sbjct:  477  KAYDLLARIHQDLLDNKGRQVDFEALDNLLERLKDVSSDKVKLV----DDILAFLAPIRH  532

Query:  409  IGQPTLPNNSLATPSPSLPINPGISHEKHEEDGYGFDANRIIAEDESGFIMSHGNHNHYF  468
              +    PN +     + +  ++ +     EDGY FD   I +++   ++  H  H+H+
Sbjct:  533  PERLGKPNAQITYTDDEIQVAK-LAGKYTTEDGYIFDPRDITSDEGDAYVTPHMTHSHWI  591

Query:  469  FKKDLTEEQIKAAQKHLEEVKTSHNGLDSLSSHEQDYPGN--AKEMKDLDKKIE--EKIA  524
              K  L+E +  AAQ + +E     GL  S+  QD GN  AK + +  +++  +K+
Sbjct:  592  KKDSLSEAERAAAQAYAKE-----KGLTPPSTDHQD-SGNTEAKGAEAIYNRVKAAKKVP  645

Query:  525  GIMKQYGVKRESIVVNKEKNAIIYPHGDHHH  555
                  Y ++       V  +   ++I PH DH+H
Sbjct:  646  LDRMPYNLQ---YTVEVKNGSLIIPHYDHYH  673


 Score = 68.2 bits (165), Expect = 8e-10
 Identities = 49/190 (25%), Positives = 80/190 (41%), Gaps = 30/190 (15%)

Query:  138  LKDAAHADNVRTKDEIN------RQKQEHVKDNEKV---NSNVAVARSQGRYTTNDGYVF  188
```

```
             LKD + +D V+  D+I      R +   K N ++   +  + VA+  G+YTT DGY+F
Sbjct: 509   LKDVS-SDKVKLVDDILAFLAPIRHPERLGKPNAQITYTDDEIQVAKLAGKYTTEDGYIF 567

Query: 189   NPADIIEDTGNAYIVPHGGHYHYIPXXXXXXXXXXXXXXXXXXXXXXNMQPSQLSYSSTASDN 248
             +P DI  D G+AY+ PH  H H+I                      + P    ST   +
Sbjct: 568   DPRDITSDEGDAYVTPHMTHSHWIKKDSLSEAERAAAQAYAKEKGLTP-----PSTDHQD 622

Query: 249   NTQSVAKGSTSKPANKSENLQSLLKELYDSPSAQRYSESDGLVFDPAKIISRTPNGVAIP 308
             +  + AKG+                 + +Y+  A +   D + ++   +    + IP
Sbjct: 623   SGNTEAKGA--------------EAIYNRVKAAKKVPLDRMPYNLQYTVEVKNGSLIIP 667

Query: 309   HGDHYHFIPY 318
             H DHYH I +
Sbjct: 668   HYDHYHNIKF 677


  Score = 63.9 bits (154), Expect = 2e-08
  Identities = 74/365 (20%), Positives = 139/365 (37%), Gaps = 85/365 (23%)

Query: 239   LSYSSTASDNNTQSVAKGSTSKPANKSENLQSLLKELYDSPSAQRYSESDGLVFDPAKII 298
             L + T +NN S  G +    K+ENL           + S+ +G+ +   +I
Sbjct: 37    LHQAQTVKENNRVSYIDGKQA--TQKTENL----------TPDEVSKREGINAEQI-VI 82

Query: 299   SRTPNGVAIPHGDHYHF----IPYSKLSALEEKIARRVPISGTGSTVSTNAKPNEVVXXX 354
               T G    HGDHYH+    +PY +  + E++ + P       +  K +++
Sbjct: 83    KITDQGYVTSHGDHYHYYNGKVPYDAI--ISEELLMKDP--------NYQLKDEDII--- 129

Query: 355   XXXXXXXXXXXXXXKELSSASDGYIFNPKDIVEETATAYI--VRHGDHFHYIPKSNQIGQP 412
                           S   GY+      V+       Y+     H D+     + N+  Q
Sbjct: 130   ----------------SEIKGGYVIK----VDGKYYVYLKDAAHADNVRTKEEINRQKQE 169

Query: 413   TLPNNSLATPSPSLPINPGISHEKHE-EDGYGFDANRIIAEDESGFIMSHGNHNHYFFKK 471
             +     TP  +    S ++  +DGY F+A+ II +    +I+ HG+H HY  K
Sbjct: 170   HSQHREGGTPRNDGAVALARSQGRYTTDDGYIFNASDIIEDTGDAYIVPHGDHYHYIPKN 229

Query: 472   DLTEEQIKAAQKHL---------------EEVKTSHNGLDSLS---------SHEQDYP 506
             +L+ ++ AA+ L                + + N + S+S         S+  +
Sbjct: 230   ELSASELAAAKAFLSGRGNLSNSRTYRRQNSDNTSRTNWVPSVSNPGTTNTNTSNNSNTN 289

Query: 507   GNAKEMKDLDKKIEEKIAGIMKQYGVKRESI------VVNKEKNAIIYPHGDHHHADPID 560
             A +  D+D +++        + Q V+ + +        + ++ N +  PHGDH+H  P
Sbjct: 290   SQASQSNDIDSLLKQLYKLPLSQRHVESDGLIFDPAQITSRTANGVAVPHGDHYHFIPYS 349

Query: 561   EHKPV 565
             +  P+
Sbjct: 350   QLSPL 354


  Score = 42.7 bits (99), Expect = 0.036
  Identities = 44/186 (23%), Positives = 63/186 (33%), Gaps = 36/186 (19%)

Query: 372   SASDGYIFNPKDIVEETATAYIVRHGDHFHYIPKSN------QIGQPTLPNNSLATPSPS 425
             +  DGYIF+P++DI +   AY+  H  H H+I K +          Q    L   PS
Sbjct: 560   TTEDGYIFDPRDITSDEGDAYVTPHMTHSHWIKKDSLSEAERAAAQAYAKEKGLTPPSTD 619

Query: 426   LPINPGISHEKHEE-------------DGYGFDANRIIAEDESGFIMSHGNHNH----- 466
             + G + K E               D   ++   +        I+ H +H H
Sbjct: 620   HQ-DSGNTEAKGAEAIYNRVKAAKKVPLDRMPYNLQYTVEVKNGSLIIPHYDHYHNIKFE 678
```

```
Query: 467  ------YFFKKDLTEEQIKAAQK----HLEEVKTSHNGLDSLSSHEQDYPGNAKEMKDLD 516
                  Y  K  +E+ A K     H E   SNG  +S H Q       +    +
Sbjct: 679  WFDEGLYEAPKGYSLEDLLATVKYYVEHPNERPHSDNGFGNASDHVQRNKNGQADTNQTE 738

Query: 517  KKIEEK 522
              K  EEK
Sbjct: 739  KPNEEK 744
```

```
 Score = 39.7 bits (91), Expect = 0.30
 Identities = 34/138 (24%), Positives = 51/138 (36%), Gaps = 19/138 (13%)

Query: 274  ELYDSPSAQRYSESDGLVFDPAKIISRTPNGVAIPHGDHYHFIPYSKLSALEEKIARRV- 332
             E+  +  A +Y+  DG +FDP  I S   +    PH  H+I   LS  E   A+
Sbjct: 549  EIQVAKLAGKYTTEDGYIFDPRDITSDEGDAYVTPHMTHSHWIKKDSLSEAERAAAQAYA 608

Query: 333  -------PISGTGSTVSTNAKPNEVVXXXXXXXXXXXXXXXXXXXKELSSASDGYIFNPKDIV 385
                    P +   +  +T AK  E +                    D  +N + V
Sbjct: 609  KEKGLTPPSTDHQDSGNTEAKGAEAIYNRVKAAK----------KVPLDRMPYNLQYTV 657

Query: 386  EETATAYIVRHGDHFHYI 403
             E    + I+ H DH+H I
Sbjct: 658  EVKNGSLIIPHYDHYHNI 675
```

tr <u>Q8DPQ2</u> **Pneumococcal histidine triad protein A [phtA] [Streptococcus 828 AA**
          pneumoniae (strain ATCC BAA-255 / R6)]

                                                                    <u>align</u>

```
 Score = 437 bits (1125), Expect = e-121
 Identities = 218/369 (59%), Positives = 271/369 (73%), Gaps = 21/369 (5%)

Query: 1   MKFSKKYIAAGSAVIVSLSLCAYALNQHRSQENKDNNRVSYVDGSQSSQKSENLTPDQVS 60
            MK +KKY+  GSA  + LS+C+Y L  ++++  K+NNRVSY+DG Q++QK+ENLTPD+VS
Sbjct: 13  MKINKKYLV-GSAAALILSVCSYELGLYQARTVKENNRVSYIDGKQATQKTENLTPDEVS 71

Query: 61  QKEGIQAEQIVIKITDQGYVTSHGDHYHYYNGKVPYDALFSEELLMKDPNYQLKDADIVN 120
            ++EGI AEQIVIKITDQGYVTSHGDHYHYYNGKVPYDA+FSEELLMKDPNY+LKD DIVN
Sbjct: 72  KREGINAEQIVIKITDQGYVTSHGDHYHYYNGKVPYDAIFSEELLMKDPNYKLKDEDIVN 131

Query: 121 EVKGGYIIKVDGKYYVYLKDAAHADNVRTKDEINRQKQEHVKDNE----KVNSNVAVARS 176
            EVKGGY+IKVDGKYYVYLKDAAHADNVRTK+EINRQKQEH + E   + + VA+ARS
Sbjct: 132 EVKGGYVIKVDGKYYVYLKDAAHADNVRTKEEINRQKQEHSQHREGGTPRNDGAVALARS 191

Query: 177 QGRYTTNDGYVFNPADIIEDTGNAYIVPHGGHYHYIPXXXXXXXXXXXXXXXXXXXXXNMQP 236
            QGRYTT+DGY+FN +DIIEDTG+AYIVPHG HYHYIP
Sbjct: 192 QGRYTTDDGYIFNASDIIEDTGDAYIVPHGDHYHYIPKNELSASELAAAEAFLSGRGNLS 251

Query: 237 SQLSYSSTASDN---------------NTQSVAKGSTSKPANKSENLQSLLKELYDSPS 280
             +  +Y   SDN                 NT +    +T+ A++S ++ SLLK+LY  P
Sbjct: 252 NSRTYRRQNSDNTSRTNWVPSVSNPGTTNTNTSNNSNTNSQASQSNDIDSLLKQLYKLPL 311

Query: 281 AQRYSESDGLVFDPAKIISRTPNGVAIPHGDHYHFIPYSKLSALEEKIARRVPISGTGST 340
            +QR+ ESDGLVFDPA+I SRT  GVA+PHGDHYHFIPYS++S LEE+IAR +P+    +
Sbjct: 312 SQRHVESDGLVFDPAQITSRTARGVAVPHGDHYHFIPYSQMSELEERIARIIPLRYRSNH 371
```

```
Query: 341 VSTNAKPNE 349
               +++P +
Sbjct: 372 WVPDSRPEQ 380
```

```
 Score = 77.8 bits (190), Expect = 1e-12
 Identities = 103/504 (20%), Positives = 189/504 (37%), Gaps = 60/504 (11%)
```

```
Query:  71 VIKITDQGYVTSHGDHYHYYNGKVPYDALFSEEL-----------LMKDPNYQLKDADI 118
           +I+ T   Y+ HGDHYHY    +P + L + EL           L     Y+ +++D
Sbjct: 208 IIEDTGDAYIVPHGDHYHY----IPKNELSASELAAAEAFLSGRGNLSNSRTYRRQNSD- 262
```

```
Query: 119 VNEVKGGYIIKVDGKYYVYLKDAAHADNVRTKDEINRQKQ-EHVKDNEKVNSNVAVARSQ 177
            N + ++ V               N T + N  Q      D + +   +
Sbjct: 263 -NTSRTNWVPSVSNP---------GTTNTNTSNNSNTNSQASQSNDIDSLLKQLYKLPLS 312
```

```
Query: 178 GRYTTNDGYVFNPADIIEDTGNAYIVPHGGHYHYIPXXXXXXXXXXXXXXXXXXXXNMQPS 237
           R+ +DG VF+PA I   T    VPHG HYH+IP                    + P
Sbjct: 313 QRHVESDGLVFDPAQITSRTARGVAVPHGDHYHFIPYSQMSELEERIA-------RIIPL 365
```

```
Query: 238 QLSYSSTASDNNTQSVAKGSTSKPANKSENLQSLLKELYDSPSAQRYSE-SDGLVFDPAK 296
           + +     D+ + +   T +P+ +   +L + S  +Q   +G VF+  K
Sbjct: 366 RYRSNHWVPDSRPEQPSPQPTPEPSPGPQPAPNLKIDSNSSLVSQLVRKVGEGYVFEE-K 424
```

```
Query: 297 IISRTPNGVAIPHGDHYHF---IPYSKLSALEEKIARRVPISGTGSTVSTNAKPNEVVXX 353
           ISR          Y F  +P  + LE K++++ +S T +   N  P +
Sbjct: 425 GISR-----------YVFAKDLPSETVKNLESKLSKQESVSHTLTAKKENVAPRDQEFY 472
```

```
Query: 354 XXXXXXXXXXXXXXXKELSSASDGYIFNPK--DIVEETATAYIVRHGDHFHYIPKSNQIGQ 411
                          E   + +    K + + + +T    D  ++          +
Sbjct: 473 DKAYNLLTEAHKALFENKGRNSDFQALDKLLERLNDESTNKEKLVDDLLAFLAPITHPER 532
```

```
Query: 412 PTLPNNSLATPSPSLPINPGISHEKHEEDGYGFDANRIIAEDESGFIMSHGNHNHYFFKK 471
             PN+ +   + I  ++ +    DGY FD + II+++  ++  H  H+H+   K
Sbjct: 533 LGKPNSQIEYTEDEVRIAQ-LADKYTTSDGYIFDEHDIISDEGDAYVTPHMGHSHWIGKD 591
```

```
Query: 472 DLTEEQIKAAQKHLEEVKTSHNGLDSLSSHEQDYPGNAKEMKDLDKKIEEKIAGIMKQYG 531
           L++++ AAQ + +E       D + + + G++        K E++I + Y
Sbjct: 592 SLSDKEKVAAQAYTKEKGILPPSPD--ADVKANPTGDSAAAIYNRVKGEKRIPLVRLPYM 649
```

```
Query: 532 VKRESIVVNKEKNAIIYPHGDHHH 555
           V+    V +    +I PH DH+H
Sbjct: 650 VEH---TVEVKNGNLIIPHKDHYH 670
```

```
 Score = 60.1 bits (144), Expect = 2e-07
 Identities = 73/372 (19%), Positives = 131/372 (34%), Gaps = 107/372 (28%)
```

```
Query: 239 LSYSSTASDNNTQSVAKGSTSKPANKSENLQSLLKELYDSPSAQRYSESDGLVFDPAKII 298
           L + T +NN S  G +    K+ENL          +   S+ +G+ +   +I
Sbjct:  38 LYQARTVKENNRVSYIDGKQA--TQKTENL----------TPDEVSKREGINAEQI-VI 83
```

```
Query: 299 SRTPNGVAIPHGDHYHF----IPYSKLSALEEKIARRVPISGTGSTVSTNAKPNEVVXXX 354
            T  G    HGDHYH+     +PY + +
Sbjct:  84 KITDQGYVTSHGDHYHYYNGKVPYDAIFS---------------------------- 112
```

```
Query: 355 XXXXXXXXXXXXXXXKELSSASDGYIFNPKDIVEETATAYIVR------------HGDHFH 401
```

```
                             +EL      Y    +DIV E     Y+++                  H D+
Sbjct: 113  ------------EELLMKDPNYKLKDEDIVNEVKGGYVIKVDGKYYVYLKDAAHADNVR 159

Query: 402  YIPKSNQIGQPTLPNNSLATPSPSLPINPGISHEKHE-EDGYGFDANRIIAEDESGFIMS 460
                 + N+  Q      +   TP      +      S   ++  +DGY F+A+ II +     +I+
Sbjct: 160  TKEEINRQKQEHSQHREGGTPRNDGAVALARSQGRYTTDDGYIFNASDIIEDTGDAYIVP 219

Query: 461  HGNHNHYFFKKDLTEEQIKAAQKHL----------------EEVKTSHNGLDSLS----- 499
                 HG+H HY  K +L+  ++ AA+  L                   +   +  N + S+S
Sbjct: 220  HGDHYHYIPKNELSASELAAAEAFLSGRGNLSNSRTYRRQNSDNTSRTNWVPSVSNPGTT 279

Query: 500  ----SHEQDYPGNAKEMKDLDKKIEEKIAGIMKQYGVKRESIV------VNKEKNAIIYP 549
                   S+  +     A +   D+D  +++       + Q V+ + +V        ++    +  P
Sbjct: 280  NTNTSNNSNTNSQASQSNDIDSLLKQLYKLPLSQRHVESDGLVFDPAQITSRTARGVAVP 339

Query: 550  HGDHHHADPIDE 561
                 HGDH+H  P +
Sbjct: 340  HGDHYHFIPYSQ 351
```

```
 Score = 57.0 bits (136), Expect = 2e-06
 Identities = 36/155 (23%), Positives = 57/155 (36%), Gaps = 20/155 (12%)

Query: 165  EKVNSNVAVARSQGRYTTNDGYVFNPADIIEDTGNAYIVPHGGHYHYIPXXXXXXXXXXX 224
                 E     V +A+   +YTT+DGY+F+  DII D G+AY+ PH GH H+I
Sbjct: 541  EYTEDEVRIAQLADKYTTSDGYIFDEHDIISDEGDAYVTPHMGHSHWIGKDSLSDKEKVA 600

Query: 225  XXXXXXXXNMQPSQLSYSSTASDNNTQSVAKGSTSKPANKSENLQSLLKELYDSPSAQRY 284
                     + P     A+     + A                +Y+     ++
Sbjct: 601  AQAYTKEKGILPPSPDADVKANPTGDSAAA-------------------IYNRVKGEKR 640

Query: 285  SESDGLVFDPAKIISRTPNGVAIPHGDHYHFIPYS 319
                   L +       +   IPH DHYH I ++
Sbjct: 641  IPLVRLPYMVEHTVEVKNGNLIIPHKDHYHNIKFA 675
```

```
tr   Q9AG74      PhpA [phpA] [Streptococcus pneumoniae]        844 AA
                                                                 align
```

```
 Score =  437 bits (1124), Expect = e-121
 Identities = 216/369 (58%), Positives = 271/369 (72%), Gaps = 21/369 (5%)

Query: 1    MKFSKKYIAAGSAVIVSLSLCAYALNQHRSQENKDNNRVSYVDGSQSSQKSENLTPDQVS 60
                 MK +KKY+  GSA + LS+C+Y L ++++  K+NNRVSY+DG Q++QK+ENLTPD+VS
Sbjct: 1    MKINKKYLV-GSAAALILSVCSYELGLYQARTVKENNRVSYIDGKQATQKTENLTPDEVS 59

Query: 61   QKEGIQAEQIVIKITDQGYVTSHGDHYHYYNGKVPYDALFSEELLMKDPNYQLKDADIVN 120
                 ++EGI AEQIVIKITDQGYVTSHGDHYHYYNGKVPYDA+ SEELLMKDPNYQLKD DI++
Sbjct: 60   KREGINAEQIVIKITDQGYVTSHGDHYHYYNGKVPYDAIISEELLMKDPNYQLKDEDIIS 119

Query: 121  EVKGGYIIKVDGKYYVYLKDAAHADNVRTKDEINRQKQEHVKDNE----KVNSNVAVARS 176
                 E+KGGY+IKVDGKYYVYLKDAAHADNVRTK+EINRQKQEH +  E    + +  VA+ARS
Sbjct: 120  EIKGGYVIKVDGKYYVYLKDAAHADNVRTKEEINRQKQEHSQHREGGTPRNDGAVALARS 179

Query: 177  QGRYTTNDGYVFNPADIIEDTGNAYIVPHGGHYHYIPXXXXXXXXXXXXXXXXXXXXNMQP 236
                 QGRYTT+DGY+FN +DIIEDTG+AYIVPHG HYHYIP
```

```
Sbjct: 180  QGRYTTDDGYIFNASDIIEDTGDAYIVPHGDHYHYIPKNELSASELAAAKAFLSGRGNLS  239

Query: 237  SQLSYSSTASDN---------------NTQSVAKGSTSKPANKSENLQSLLKELYDSPS  280
            +  +Y   SDN              NT +    +T+ A++S ++ SLLK+LY   P
Sbjct: 240  NSRTYRRQNSDNTSRTNWVPSVSNPGTTNTNTSNNSNTNSQASQSNDIDSLLKQLYKLPL  299

Query: 281  AQRYSESDGLVFDPAKIISRTPNGVAIPHGDHYHFIPYSKLSALEEKIARRVPISGTGST  340
            +QR+ ESDGL+FDPA+I SRT NGVA+PHGDHYHFIPYS+LS LEEK+AR +P+    +
Sbjct: 300  SQRHVESDGLIFDPAQITSRTANGVAVPHGDHYHFIPYSQLSPLEEKLARIIPLRYRSNH  359

Query: 341  VSTNAKPNE  349
            +++P +
Sbjct: 360  WVPDSRPEQ  368


  Score = 84.0 bits (206), Expect = 1e-14
  Identities = 111/511 (21%), Positives = 190/511 (36%), Gaps = 70/511 (13%)

Query: 71   VIKITDQGYVTSHGDHYHYYNGKVPYDALFSEEL-----------LMKDPNYQLKDADI  118
            +I+ T  Y+  HGDHYHY    +P + L + EL           L     Y+ +++D
Sbjct: 196  IIEDTGDAYIVPHGDHYHY----IPKNELSASELAAAKAFLSGRGNLSNSRTYRRQNSD-  250

Query: 119  VNEVKGGYIIKVDGKYYVYLKDAAHADNVRTKDEINRQKQ-EHVKDNEKVNSNVAVARSQ  177
            N + ++ V        N  T + N  Q    D + +   +
Sbjct: 251  -NTSRTNWVPSVSNP---------GTTNTNTSNNSNTNSQASQSNDIDSLLKQLYKLPLS  300

Query: 178  GRYTTNDGYVFNPADIIEDTGNAYIVPHGGHYHYIPXXXXXXXXXXXXXXXXXXXX-----  232
            R+ +DG +F+PA I   T N  VPHG HYH+IP
Sbjct: 301  QRHVESDGLIFDPAQITSRTANGVAVPHGDHYHFIPYSQLSPLEEKLARIIPLRYRSNHW  360

Query: 233  --NMQPSQLSYSSTASDNNTQSVAKGSTSKPANKSENLQSLLKELYDSPSAQRYSESDGL  290
              + +P Q S  ST   + +    A        P+N  +  + L+KE         E +G+
Sbjct: 361  VPDSRPEQPSPQSTPEPSPSPQPAPNPQPAPSNPID--EKLVKEAVRKVGDGYVFEENGV  418

Query: 291  V-FDPAKIIS-RTPNGVAIPHGDHYHFIPYSKLSALEEKIARRVPISGTGSTVSTNAKPN  348
            + + PAK +S  T  G+          SKL A +E ++ ++     T   S    N
Sbjct: 419  PRYIPAKDLSAETAAGID-----------SKL-AKQESLSHKLGAKKTDLPSSDREFYN  465

Query: 349  EVVXXXXXXXXXXXXXXXXXXKELSSASDGYIFNPKDIVEETATAYIVRHGDHFHYIPKSNQ  408
            +                  + A D +   KD+ +      D  ++
Sbjct: 466  KAYDLLARIHQDLLDNKGRQVDFEALDNLLERLKDVSSDKVKLV----DDILAFLAPIRH  521

Query: 409  IGQPTLPNNSLATPSPSLPINPGISHEKHEEDGYGFDANRIIAEDESGFIMSHGNHNHYF  468
            +    PN +          + +  ++ +  EDGY FD   I +++    ++  H  H+H+
Sbjct: 522  PERLGKPNAQITYTDDEIQVAK-LAGKYTTEDGYIFDPRDITSDEGDAYVTPHMTHSHWI  580

Query: 469  FKKDLTEEQIKAAQKHLEEVKTSHNGLDSLSSHEQDYPGN--AKEMKDLDKKIE--EKIA  524
            K  L+E + AAQ ++ +E       GL   S+ QD  GN  AK + +  +++   +K+
Sbjct: 581  KKDSLSEAERAAAQAYAKE-----KGLTPPSTDHQD-SGNTEAKGAEAIYNRVKAAKKVP  634

Query: 525  GIMKQYGVKRESIVVNKEKNAIIYPHGDHHH  555
            + ++     V + ++I PH DH+H
Sbjct: 635  LDRMPYNLQ---YTVEVKNGSLIIPHYDHYH  662


  Score = 68.2 bits (165), Expect = 8e-10
  Identities = 49/190 (25%), Positives = 80/190 (41%), Gaps = 30/190 (15%)
```

```
Query: 138 LKDAAHADNVRTKDEIN------RQKQEHVKDNEKV---NSNVAVARSQGRYTTNDGYVF 188
            LKD + +D V+  D+I       R +   K N ++  +  + VA+  G+YTT DGY+F
Sbjct: 498 LKDVS-SDKVKLVDDILAFLAPIRHPERLGKPNAQITYTDDEIQVAKLAGKYTTEDGYIF 556

Query: 189 NPADIIEDTGNAYIVPHGGHYHYIPXXXXXXXXXXXXXXXXXXXXXNMQPSQLSYSSTASDN 248
            +P DI  D G+AY+ PH  H H+I                      + P     ST  +
Sbjct: 557 DPRDITSDEGDAYVTPHMTHSHWIKKDSLSEAERAAAQAYAKEKGLTP-----PSTDHQD 611

Query: 249 NTQSVAKGSTSKPANKSENLQSLLKELYDSPSAQRYSESDGLVFDPAKIISRTPNGVAIP 308
            +  + AKG+                + +Y+  A +   D + ++  +     + IP
Sbjct: 612 SGNTEAKGA--------------EAIYNRVKAAKKVPLDRMPYNLQYTVEVKNGSLIIP 656

Query: 309 HGDHYHFIPY 318
            H DHYH I +
Sbjct: 657 HYDHYHNIKF 666


 Score = 63.2 bits (152), Expect = 3e-08
 Identities = 74/365 (20%), Positives = 139/365 (37%), Gaps = 85/365 (23%)

Query: 239 LSYSSTASDNNTQSVAKGSTSKPANKSENLQSLLKELYDSPSAQRYSESDGLVFDPAKII 298
            L + T +NN S  G +   K+ENL         +   S+ +G+  +   +I
Sbjct: 26  LYQARTVKENNRVSYIDGKQA--TQKTENL-----------TPDEVSKREGINAEQI-VI 71

Query: 299 SRTPNGVAIPHGDHYHF----IPYSKLSALEEKIARRVPISGTGSTVSTNAKPNEVVXXX 354
              T G    HGDHYH+      +PY + + E++ + P      + K  +++
Sbjct: 72  KITDQGYVTSHGDHYHYYNGKVPYDAI--ISEELLMKDP--------NYQLKDEDII--- 118

Query: 355 XXXXXXXXXXXXXKELSSASDGYIFNPKDIVEETATAYI--VRHGDHFHYIPKSNQIGQP 412
                         S   GY+      V+      Y+    H D+    + N+   Q
Sbjct: 119 ---------------SEIKGGYVIK----VDGKYYVYLKDAAHADNVRTKEEINRQKQE 158

Query: 413 TLPNNSLATPSPSLPINPGISHEKHE-EDGYGFDANRIIAEDESGFIMSHGNHNHYFFKK 471
             +    TP   +    S  ++ +DGY F+A+ II +     +I+ HG+H HY   K
Sbjct: 159 HSQHREGGTPRNDGAVALARSQGRYTTDDGYIFNASDIIEDTGDAYIVPHGDHYHYIPKN 218

Query: 472 DLTEEQIKAAQKHL---------------EEVKTSHNGLDSLS---------SHEQDYP 506
            +L+  ++ AA+  L                + + N + S+S         S+   +
Sbjct: 219 ELSASELAAAKAFLSGRGNLSNSRTYRRQNSDNTSRTNWVPSVSNPGTTNTNTSNNSNTN 278

Query: 507 GNAKEMKDLDKKIEEKIAGIMKQYGVKRESI------VVNKEKNAIIYPHGDHHHADPID 560
            A + D+D +++     + Q V+ + +       + ++ N +  PHGDH+H  P
Sbjct: 279 SQASQSNDIDSLLKQLYKLPLSQRHVESDGLIFDPAQITSRTANGVAVPHGDHYHFIPYS 338

Query: 561 EHKPV 565
            +  P+
Sbjct: 339 QLSPL 343


 Score = 42.7 bits (99), Expect = 0.036
 Identities = 44/186 (23%), Positives = 63/186 (33%), Gaps = 36/186 (19%)

Query: 372 SASDGYIFNPKDIVEETATAYIVRHGDHFHYIPKSN------QIGQPTLPNNSLATPSPS 425
            +  DGYIF+P+DI  +   AY+  H  H H+I K +      Q     L    PS
Sbjct: 549 TTEDGYIFDPRDITSDEGDAYVTPHMTHSHWIKKDSLSEAERAAAQAYAKEKGLTPPSTD 608

Query: 426 LPINPGISHEKHEE--------------DGYGFDANRIIAEDESGFIMSHGNHNH----- 466
             + G + K E               D  ++   +       I+ H +H H
```

```
Sbjct:  609  HQ-DSGNTEAKGAEAIYNRVKAAKKVPLDRMPYNLQYTVEVKNGSLIIPHYDHYHNIKFE 667

Query:  467  ------YFFKKDLTEEQIKAAQK----HLEEVKTSHNGLDSLSSHEQDYPGNAKEMKDLD 516
                    Y K +E+A K     H E  S NG + S H Q        +   +
Sbjct:  668  WFDEGLYEAPKGYSLEDLLATVKYYVEHPNERPHSDNGFGNASDHVQRNKNGQADTNQTE 727

Query:  517  KKIEEK 522
                K  EEK
Sbjct:  728  KPNEEK 733
```

Score = 39.7 bits (91), Expect = 0.30
Identities = 34/138 (24%), Positives = 51/138 (36%), Gaps = 19/138 (13%)

```
Query:  274  ELYDSPSAQRYSESDGLVFDPAKIISRTPNGVAIPHGDHYHFIPYSKLSALEEKIARRV- 332
             E+  +  A +Y+  DG +FDP  I S   +    PH  H H+I    LS  E    A+
Sbjct:  538  EIQVAKLAGKYTTEDGYIFDPRDITSDEGDAYVTPHMTHSHWIKKDSLSEAERAAAQAYA 597

Query:  333  -------PISGTGSTVSTNAKPNEVVXXXXXXXXXXXXXXXXXXXKELSSASDGYIFNPKDIV 385
                    P +     + +T AK  E +                    D  +N + V
Sbjct:  598  KEKGLTPPSTDHQDSGNTEAKGAEAIYNRVKAAK----------KVPLDRMPYNLQYTV 646

Query:  386  EETATAYIVRHGDHFHYI 403
             E    + I+ H DH+H I
Sbjct:  647  EVKNGSLIIPHYDHYHNI 664
```

tr Q9AHT9 **Pneumococcal histidine triad A protein [phtA] [Streptococcus** 816 AA
             pneumoniae]

                                                                                 align

Score = 435 bits (1119), Expect = e-120
Identities = 217/369 (58%), Positives = 270/369 (72%), Gaps = 21/369 (5%)

```
Query:  1    MKFSKKYIAAGSAVIVSLSLCAYALNQHRSQENKDNNRVSYVDGSQSSQKSENLTPDQVS 60
             MK +KKY+ GSA  + LS+C+Y L  ++++  K+NNRVSY+DG Q++QK+ENLTPD+VS
Sbjct:  1    MKINKKYLV-GSAAALILSVCSYELGLYQARTVKENNRVSYIDGKQATQKTENLTPDEVS 59

Query:  61   QKEGIQAEQIVIKITDQGYVTSHGDHYHYYNGKVPYDALFSEELLMKDPNYQLKDADIVN 120
             ++EGI AEQIVIKITDQGYVTSHGDHYHYYNGKVPYDA+ SEELLMKDPNY+LKD DIVN
Sbjct:  60   KREGINAEQIVIKITDQGYVTSHGDHYHYYNGKVPYDAIISEELLMKDPNYKLKDEDIVN 119

Query:  121  EVKGGYIIKVDGKYYVYLKDAAHADNVRTKDEINRQKQEHVKDNE----KVNSNVAVARS 176
             EVKGGY+IKVDGKYYVYLKDAAHADNVRTK+EINRQKQEH + E   + + VA+ARS
Sbjct:  120  EVKGGYVIKVDGKYYVYLKDAAHADNVRTKEEINRQKQEHSQHREGGTPRNDGAVALARS 179

Query:  177  QGRYTTNDGYVFNPADIIEDTGNAYIVPHGGHYHYIPXXXXXXXXXXXXXXXXXXXNMQP 236
             QGRYTT+DGY+FN +DIIEDTG+AYIVPHG HYHYIP
Sbjct:  180  QGRYTTDDGYIFNASDIIEDTGDAYIVPHGDHYHYIPKNELSASELAAAEAFLSGRGNLS 239

Query:  237  SQLSYSSTASDN---------------NTQSVAKGSTSKPANKSENLQSLLKELYDSPS 280
             +  +Y   SDN               NT +  +T+  A++S ++ SLLK+LY  P
Sbjct:  240  NSRTYRRQNSDNTSRTNWVPSVSNPGTTNTNTSNNSNTNSQASQSNDIDSLLKQLYKLPL 299

Query:  281  AQRYSESDGLVFDPAKIISRTPNGVAIPHGDHYHFIPYSKLSALEEKIARRVPISGTGST 340
             +QR+ ESDGLVFDPA+I SRT  GVA+PHGDHYHFIPYS++S LEE+IAR +P+     +
```

```
Sbjct: 300 SQRHVESDGLVFDPAQITSRTARGVAVPHGDHYHFIPYSQMSELEERIARIIPLRYRSNH 359

Query: 341 VSTNAKPNE 349
               +++P +
Sbjct: 360 WVPDSRPEQ 368


 Score = 75.5 bits (184), Expect = 5e-12
 Identities = 102/504 (20%), Positives = 188/504 (37%), Gaps = 60/504 (11%)

Query: 71  VIKITDQGYVTSHGDHYHYYNGKVPYDALFSEEL-----------LMKDPNYQLKDADI 118
           +I+ T Y+ HGDHYHY    +P + L + EL          L    Y+ +++D
Sbjct: 196 IIEDTGDAYIVPHGDHYHY----IPKNELSASELAAAEAFLSGRGNLSNSRTYRRQNSD- 250

Query: 119 VNEVKGGYIIKVDGKYYVYLKDAAHADNVRTKDEINRQKQ-EHVKDNEKVNSNVAVARSQ 177
           N + ++ V            N T + N Q     D + +
Sbjct: 251 -NTSRTNWVPSVSNP--------GTTNTNTSNNSNTNSQASQSNDIDSLLKQLYKLPLS 300

Query: 178 GRYTTNDGYVFNPADIIEDTGNAYIVPHGGHYHYIPXXXXXXXXXXXXXXXXXXXXNMQPS 237
           R+ +DG VF+PA I   T      VPHG HYH+IP                    + P
Sbjct: 301 QRHVESDGLVFDPAQITSRTARGVAVPHGDHYHFIPYSQMSELEERIA-------RIIPL 353

Query: 238 QLSYSSTASDNNTQSVAKGSTSKPANKSENLQSLLKELYDSPSAQRYSE-SDGLVFDPAK 296
           + + D+ + + T +P+ + +L + S +Q + +G VF+ K
Sbjct: 354 RYRSNHWVPDSRPEQPSPQPTPEPSPGPQPAPNLKIDSNSSLVSQLVRKVGEGYVFEE-K 412

Query: 297 IISRTPNGVAIPHGDHYHF---IPYSKLSALEEKIARRVPISGTGSTVSTNAKPNEVVXX 353
           ISR        Y F +P + LE K++++ +S T + N P +
Sbjct: 413 GISR-----------YVFAKDLPSETVKNLESKLSKQESVSHTLTAKKENVAPRDQEFY 460

Query: 354 XXXXXXXXXXXXXXXKELSSASDGYIFNPK--DIVEETATAYIVRHGDHFHYIPKSNQIGQ 411
                          + +     K + + + +T        D ++         +
Sbjct: 461 DKAYNLLTEAHKALFXNKGRNSDFQALDKLLERLNDESTNKEKLVDDLLAFLAPITHPER 520

Query: 412 PTLPNNSLATPSPSLPINPGISHEKHEEDGYGFDANRIIAEDESGFIMSHGNHNHYFFKK 471
           PN+ +  + I ++ + DGY FD + II+++ ++ H H+H+ K
Sbjct: 521 LGKPNSQIEYTEDEVRIAQ-LADKYTTSDGYIFDEHDIISDEGDAYVTPHMGHSHWIGKD 579

Query: 472 DLTEEQIKAAQKHLEEVKTSHNGLDSLSSHEQDYPGNAKEMKDLDKKIEEKIAGIMKQYG 531
           L++++ AAQ + +E     D + + + G++      K E++I + Y
Sbjct: 580 SLSDKEKVAAQAYTKEKGILPPSPD--ADVKANPTGDSAAAIYNRVKGEKRIPLVRLPYM 637

Query: 532 VKRESIVVNKEKNAIIYPHGDHHH 555
           V+    V +   +I PH DH+H
Sbjct: 638 VEH---TVEVKNGNLIIPHKDHYH 658


 Score = 60.1 bits (144), Expect = 2e-07
 Identities = 73/372 (19%), Positives = 131/372 (34%), Gaps = 107/372 (28%)

Query: 239 LSYSSTASDNNTQSVAKGSTSKPANKSENLQSLLKELYDSPSAQRYSESDGLVFDPAKII 298
           L + T +NN S   G +     K+ENL           +   S+ +G+ +   +I
Sbjct: 26  LYQARTVKENNRVSYIDGKQA--TQKTENL-----------TPDEVSKREGINAEQI-VI 71

Query: 299 SRTPNGVAIPHGDHYHF----IPYSKLSALEEKIARRVPISGTGSTVSTNAKPNEVVXXX 354
           T G   HGDHYH+    +PY + +
Sbjct: 72  KITDQGYVTSHGDHYHYYNGKVPYDAIIS---------------------------- 100
```

```
Query: 355 XXXXXXXXXXXXXKELSSASDGYIFNPKDIVEETATAYIVR-----------HGDHFH 401
                        +EL       Y   +DIV E      Y+++              H D+
Sbjct: 101 -----------EELLMKDPNYKLKDEDIVNEVKGGYVIKVDGKYYVYLKDAAHADNVR 147

Query: 402 YIPKSNQIGQPTLPNNSLATPSPSLPINPGISHEKHE-EDGYGFDANRIIAEDESGFIMS 460
                 + N+ Q      +    TP    +    S  ++ +DGY F+A+ II +    +I+
Sbjct: 148 TKEEINRQKQEHSQHREGGTPRNDGAVALARSQGRYTTDDGYIFNASDIIEDTGDAYIVP 207

Query: 461 HGNHNHYFFKKDLTEEQIKAAQKHL----------------EEVKTSHNGLDSLS----- 499
                HG+H HY  K +L+  ++ AA+   L               + + N + S+S
Sbjct: 208 HGDHYHYIPKNELSASELAAAEAFLSGRGNLSNSRTYRRQNSDNTSRTNWVPSVSNPGTT 267

Query: 500 ----SHEQDYPGNAKEMKDLDKKIEEKIAGIMKQYGVKRESIV------VNKEKNAIIYP 549
                   S+ +    A + D+D +++        + Q V+ + +V      ++    + P
Sbjct: 268 NTNTSNNSNTNSQASQSNDIDSLLKQLYKLPLSQRHVESDGLVFDPAQITSRTARGVAVP 327

Query: 550 HGDHHHADPIDE 561
                HGDH+H  P +
Sbjct: 328 HGDHYHFIPYSQ 339


  Score = 57.0 bits (136), Expect = 2e-06
  Identities = 36/155 (23%), Positives = 57/155 (36%), Gaps = 20/155 (12%)

Query: 165 EKVNSNVAVARSQGRYTTNDGYVFNPADIIEDTGNAYIVPHGGHYHYIPXXXXXXXXXXX 224
                E     V +A+    +YTT+DGY+F+  DII D G+AY+ PH GH H+I
Sbjct: 529 EYTEDEVRIAQLADKYTTSDGYIFDEHDIISDEGDAYVTPHMGHSHWIGKDSLSDKEKVA 588

Query: 225 XXXXXXXXNMQPSQLSYSSTASDNNTQSVAKGSTSKPANKSENLQSLLKELYDSPSAQRY 284
                   + P     A+    + A               +Y+    ++
Sbjct: 589 AQAYTKEKGILPPSPDADVKANPTGDSAAA--------------------IYNRVKGEKR 628

Query: 285 SESDGLVFDPAKIISRTPNGVAIPHGDHYHFIPYS 319
                L +    +     + IPH DHYH I ++
Sbjct: 629 IPLVRLPYMVEHTVEVKNGNLIIPHKDHYHNIKFA 663




tr Q8DQ08  Pneumococcal histidine triad protein D [phtD] [Streptococcus  853 AA
           pneumoniae (strain ATCC BAA-255 / R6)]
                                                                    align


  Score =  434 bits (1117), Expect = e-120
  Identities = 218/357 (61%), Positives = 272/357 (76%), Gaps = 11/357 (3%)

Query: 1   MKFSKKYIAAGSAVIVSLSLCAYALNQHRS-QENKDNNRVSYVDGSQSSQKSENLTPDQV 59
               MK +KKY+A GS  +++LS+C+Y L +H++ Q  K++NRVSY+DG Q+ QK+ENLTPD+V
Sbjct: 1   MKINKKYLA-GSVAVLALSVCSYELGRHQAGQVKKESNRVSYIDGDQAGQKAENLTPDEV 59

Query: 60  SQKEGIQAEQIVIKITDQGYVTSHGDHYHYYNGKVPYDALFSEELLMKDPNYQLKDADIV 119
               S++EGI AEQIVIKITDQGYVTSHGDHYHYYNGKVPYDA+ SEELLMKDPNYQLKD+DIV
Sbjct: 60  SKREGINAEQIVIKITDQGYVTSHGDHYHYYNGKVPYDAIISEELLMKDPNYQLKDSDIV 119

Query: 120 NEVKGGYIIKVDGKYYVYLKDAAHADNVRTKDEINRQKQEHVKD-NEKVNSNVAVARSQG 178
               NE+KGGY+IKVDGKYYVYLKDAAHADN+RTK+EI RQKQE   + N + ++ VA AR+QG
Sbjct: 120 NEIKGGYVIKVDGKYYVYLKDAAHADNIRTKEEIKRQKQERSHNHNSRADNAVAAARAQG 179
```

```
Query:  179  RYTTNDGYVFNPADIIEDTGNAYIVPHGGHYHYIPXXXXXXXXXXXXXXXXXXXXXXNMQPSQ  238
             RYTT+DGY+FN +DIIEDTG+AYIVPHG HYHYIP                         Q S+
Sbjct:  180  RYTTDDGYIFNASDIIEDTGDAYIVPHGDHYHYIP--KSDLSASELAAAQAYWNGKQGSR    237

Query:  239  LSYSSTASDNNTQ---SVAKGSTSKPA---NKSENLQSLLKELYDSPSAQRYSESDGLVF   292
             S SS+ + N Q    S     T  P    N+ EN+ SLL+ELY  P ++R+ ESDGL+F
Sbjct:  238  PSSSSSHNANPAQPRLSENHNLTVTPTYHQNQGENISSLLRELYAKPLSERHVESDGLIF   297

Query:  293  DPAKIISRTPNGVAIPHGDHYHFIPYSKLSALEEKIARRVPISGTGSTVSTNAKPNE     349
             DPA+I SRT NGVA+PHGDHYHFIPYS+LS LEEK+AR +P+     +   +++P +
Sbjct:  298  DPAQITSRTANGVAVPHGDHYHFIPYSQLSPLEEKLARIIPLRYRSNHWVPDSRPEQ     354


 Score = 84.3 bits (207), Expect = 1e-14
 Identities = 90/385 (23%), Positives = 134/385 (34%), Gaps = 73/385 (18%)

Query:  193  IIEDTGNAYIVPHGGHYHYI-----------------PXXXXXXXXXXXXXXXXXXXXNM  234
             +I+ T  Y+  HG HYHY                  P                     +
Sbjct:   71  VIKITDQGYVTSHGDHYHYYNGKVPYDAIISEELLMKDPNYQLKDSDIVNEIKGGYVIKV  130

Query:  235  QPSQLSYSSTASDNNTQSVAKGSTSKPANKSENLQSLLKELYDSPSAQ-RYSESDGLVFD  293
                   Y  A+ +    +     +     +S N S    + AQ RY+  DG +F+
Sbjct:  131  DGKYYVYLKDAAHADNIRTKEEIKRQKQERSHNHNSRADNAVAAARAQGRYTTDDGYIFN  190

Query:  294  PAKIISRTPNGVAIPHGDHYHFIPYSKLSALEEKIARRVPISGTGSTVST----NAKP--  347
             + II T +  +PHGDHYH+IP S LSA E   A+       GS S+    NA P
Sbjct:  191  ASDIIEDTGDAYIVPHGDHYHYIPKSDLSASELAAAQAYWNGKQGSRPSSSSSHNANPAQ  250

Query:  348  -------NEVVXXXXXXXXXXXXXXXXKELSSA--------SDGYIFNPKDIVEETATAY  392
                    N  V                 +EL +        SDG IF+P  I   TA
Sbjct:  251  PRLSENHNLTVTPTYHQNQGENISSLLRELYAKPLSERHVESDGLIFDPAQITSRTANGV  310

Query:  393  IVRHGDHFHYIPKS------------------------NQIGQPTLPNNSLATPSPSLP  427
             V HGDH+H+IP S                        ++ QP+ +    +PSP
Sbjct:  311  AVPHGDHYHFIPYSQLSPLEEKLARIIPLRYRSNHWVPDSRPEQPSPQSTPEPSPSPQPA  370

Query:  428  INPGISHEKHEEDGYGFDANRIIAEDESGFIMSHGNHNHYFFKKDLTEEQIKAAQKHL-E  486
             NP +   ++   +A R + +  G++        Y  KDL+ E        L +
Sbjct:  371  PNPQPAPSNPIDEKLVKEAVRKVGD---GYVFEENGVPRYIPAKDLSAETAAGIDSKLAK  427

Query:  487  EVKTSHNGLDSLSSHEQDYPGNAKE  511
             +   SH    L + + D P + +E
Sbjct:  428  QESLSH----KLGAKKTDLPSSDRE  448


 Score = 68.2 bits (165), Expect = 8e-10
 Identities = 49/190 (25%), Positives = 80/190 (41%), Gaps = 30/190 (15%)

Query:  138  LKDAAHADNVRTKDEIN------RQKQEHVKDNEKV---NSNVAVARSQGRYTTNDGYVF  188
             LKD + +D V+ D+I       R +  K N ++   +  + VA+  G+YTT DGY+F
Sbjct:  484  LKDVS-SDKVKLVDDILAFLAPIRHPERLGKPNAQITYTDDEIQVAKLAGKYTTEDGYIF  542

Query:  189  NPADIIEDTGNAYIVPHGGHYHYIPXXXXXXXXXXXXXXXXXXXXXXXNMQPSQLSYSSTASDN  248
             +P DI  D G+AY+ PH  H H+I                       + P    ST  +
Sbjct:  543  DPRDITSDEGDAYVTPHMTHSHWIKKDSLSEAERAAAQAYAKEKGLTP-----PSTDHQD     597

Query:  249  NTQSVAKGSTSKPANKSENLQSLLKELYDSPSAQRYSESDGLVFDPAKIISRTPNGVAIP  308
             + + AKG+          + +Y+   A +     D + ++    +        + IP
```

```
Sbjct: 598 SGNTEAKGA--------------EAIYNRVKAAKKVPLDRMPYNLQYTVEVKNGSLIIP 642

Query: 309 HGDHYHFIPY 318
           H DHYH I +
Sbjct: 643 HYDHYHNIKF 652
```

```
 Score = 58.2 bits (139), Expect = 8e-07
 Identities = 60/306 (19%), Positives = 112/306 (35%), Gaps = 84/306 (27%)

Query: 297 IISRTPNGVAIPHGDHYHF----IPYSKLSALEEKIARRVPISGTGSTVSTNAKPNEVVX 352
           +I   T  G    HGDHYH+    +PY  + +
Sbjct: 71  VIKITDQGYVTSHGDHYHYYNGKVPYDAIIS--------------------------- 101

Query: 353 XXXXXXXXXXXXXXXXXKELSSASDGYIFNPKDIVEETATAYIVR-HGDHFHYIPKS----- 406
                            +EL     Y    DIV E     Y+++  G ++ Y+   +
Sbjct: 102 ---------------EELLMKDPNYQLKDSDIVNEIKGGYVIKVDGKYYVYLKDAAHADN 146

Query: 407 --------NQIGQPTLPNNSLATPSPSLPINPGISHEKHEEDGYGFDANRIIAEDESGFI 458
                   Q + +  +NS A + +     G       +DGY F+A+ II +    +I
Sbjct: 147 IRTKEEIKRQKQERSHNHNSRADNAVAAARAQG---RYTTDDGYIFNASDIIEDTGDAYI 203

Query: 459 MSHGNHNHYFFKKDLTEEQIKAAQKHLE-------EVKTSHNGLDSLSSHEQDY------ 505
           + HG+H HY  K DL+  ++ AAQ +         +SHN    +      +++
Sbjct: 204 VPHGDHYHYIPKSDLSASELAAAQAYWNGKQGSRPSSSSSHNANPAQPRLSENHNLTVTP 263

Query: 506 PGNAKEMKDLDKKIEEKIAGIMKQYGVKRESI------VVNKEKNAIIYPHGDHHHADPI 559
              +   + +++   + E  A + + V+ + +       + ++ N +  PHGDH+H  P
Sbjct: 264 TYHQNQGENISSLLRELYAKPLSERHVESDGLIFDPAQITSRTANGVAVPHGDHYHFIPY 323

Query: 560 DEHKPV 565
            +  P+
Sbjct: 324 SQLSPL 329
```

```
 Score = 47.4 bits (111), Expect = 0.001
 Identities = 61/260 (23%), Positives = 90/260 (34%), Gaps = 51/260 (19%)

Query: 372 SASDGYIFNPKDIVEETATAYIVRHGDHFHYIPKSN------QIGQPTLPNNSLATPSPS 425
           +  DGYIF+P+DI  +    AY+  H  H H+I K +        Q      L  PS
Sbjct: 535 TTEDGYIFDPRDITSDEGDAYVTPHMTHSHWIKKDSLSEAERAAAQAYAKEKGLTPPSTD 594

Query: 426 LPINPGISHEKHEE-------------DGYGFDANRIIAEDESGFIMSHGNHNH----- 466
           + G + K E             D ++   +       I+ H +H H
Sbjct: 595 HQ-DSGNTEAKGAEAIYNRVKAAKKVPLDRMPYNLQYTVEVKNGSLIIPHYDHYHNIKFE 653

Query: 467 ------YFFKKDLTEEQIKAAQK----HLEEVKTSHNGLDSLSSHEQDYPGNAKEMKDLD 516
                 Y K   + E + A K    H E   S NG  + S H Q        +     +
Sbjct: 654 WFDEGLYEAPKGYSLEDLLATVKYYVEHPNERPHSDNGFGNASDHVQRNKNGQADTNQTE 713

Query: 517 KKIEEKIAGIMKQYGVKRESIVVNKEKNAIIYPHGDHHHADPIDEHKPVGIGHSHSNYEL 576
           K  EEK         + E     +KE + +   P  H +D + H VG+  S  N   L
Sbjct: 714 KPNEEK--------PQTEKPEEDKEHDEVSEP--THPESDEKENH--VGLNPSADN--L 758

Query: 577 FKPEEGVAKKEGNKVYTGEE 596
           +KP     + E    T +E
Sbjct: 759 YKPSTDTEETEEEAEDTTDE 778
```

```
Score = 39.7 bits (91), Expect = 0.30
Identities = 34/138 (24%), Positives = 51/138 (36%), Gaps = 19/138 (13%)

Query: 274 ELYDSPSAQRYSESDGLVFDPAKIISRTPNGVAIPHGDHYHFIPYSKLSALEEKIARRV- 332
            E+ + A +Y+ DG +FDP I S +   PH H H+I  LS E   A+
Sbjct: 524 EIQVAKLAGKYTTEDGYIFDPRDITSDEGDAYVTPHMTHSHWIKKDSLSEAERAAAQAYA 583

Query: 333 -------PISGTGSTVSTNAKPNEVVXXXXXXXXXXXXXXXXXXKELSSASDGYIFNPKDIV 385
                   P +    + +T AK  E +                      D  +N + V
Sbjct: 584 KEKGLTPPSTDHQDSGNTEAKGAEAIYNRVKAAK----------KVPLDRMPYNLQYTV 632

Query: 386 EETATAYIVRHGDHFHYI 403
            E    + I+ H DH+H I
Sbjct: 633 EVKNGSLIIPHYDHYHNI 650
```

tr   Q6T8D7   **Pneumococcal protein BVH-11-3 [Streptococcus**   817 AA
                  pneumoniae]   align

```
Score =  431 bits (1109), Expect = e-119
Identities = 217/355 (61%), Positives = 266/355 (74%), Gaps = 22/355 (6%)

Query: 1   MKFSKKYIAAGSAVIVSLSLCAYALNQHRSQENKDNNRVSYVDGSQSSQKSENLTPDQVS 60
            MK +KKY+  GSA + LS+C+Y L ++++  K+NNRVSY+DG Q++QK+ENLTPD+VS
Sbjct: 1   MKINKKYLV-GSAAALILSVCSYELGLYQARTVKENNRVSYIDGKQATQKTENLTPDEVS 59

Query: 61  QKEGIQAEQIVIKITDQGYVTSHGDHYHYYNGKVPYDALFSEELLMKDPNYQLKDADIVN 120
            ++EGI AEQIVIKITDQGYVTSHGDHYHYYNGKVPYDA+FSEELLMKDPNY+LKD DIVN
Sbjct: 60  KREGINAEQIVIKITDQGYVTSHGDHYHYYNGKVPYDAIFSEELLMKDPNYKLKDEDIVN 119

Query: 121 EVKGGYIIKVDGKYYVYLKDAAHADNVRTKDEINRQKQEHVKDNE----KVNSNVAVARS 176
            EVKGGY+IKVDGKYYVYLKDAAHADNVRTK+EINRQKQEH +  E    + + VA+ARS
Sbjct: 120 EVKGGYVIKVDGKYYVYLKDAAHADNVRTKEEINRQKQEHSQHREGGTPRNDGAVALARS 179

Query: 177 QGRYTTNDGYVFNPADIIEDTGNAYIVPHGGHYHYIPXXXXXXXXXXXXXXXXXXXXNMQP 236
            QGRYTT+DGY+FN +DIIEDTG+AYIVPHG HYHYIP
Sbjct: 180 QGRYTTDDGYIFNASDIIEDTGDAYIVPHGDHYHYIPKNELSASELAAAEAFLSGRGNLS 239

Query: 237 SQLSYSSTASDN----------------NTQSVAKGSTSKPANKS-ENLQSLLKELYDSP 279
            + +Y   SDN                NT +    +T+ A++S E++ SLLK+LY  P
Sbjct: 240 NSRTYRRQNSDNTSRTNWVPSVSNPGTTNTNTSNNSNTNGQASQSNEDVDSLLKQLYALP 299

Query: 280 SAQRYSESDGLVFDPAKIISRTPNGVAIPHGDHYHFIPYSKLSALEEKIARRVPI 334
            ++R+ ESDGLVFDPA+I SRT  GVA+PHGDHYHFIPYS++S LEE+IAR +P+
Sbjct: 300 LSKRHVESDGLVFDPAQITSRTARGVAVPHGDHYHFIPYSQMSELEERIARIIPL 354
```

```
Score = 79.7 bits (195), Expect = 3e-13
Identities = 102/503 (20%), Positives = 189/503 (37%), Gaps = 57/503 (11%)

Query: 71  VIKITDQGYVTSHGDHYHYYNGKVPYDALFSEEL-----------LMKDPNYQLKDADI 118
            +I+ T  Y+ HGDHYHY    +P + L + EL           L   Y+ +++D
Sbjct: 196 IIEDTGDAYIVPHGDHYHY----IPKNELSASELAAAEAFLSGRGNLSNSRTYRRQNSD- 250

Query: 119 VNEVKGGYIIKVDGKYYVYLKDAAHADNVRTKDEINRQKQEHVKDNEKVNSNVAVARSQG 178
```

```
                N   +  ++  V                    N           N Q   +  +D + +    +
Sbjct: 251  -NTSRTNWVPSVSNP-------GTTNTNTSNNSNTNGQASQSNEDVDSLLKQLYALPLSK 302

Query: 179  RYTTNDGYVFNPADIIEDTGNAYIVPHGGHYHYIPXXXXXXXXXXXXXXXXXXXXXNMQPSQ 238
            R+   +DG VF+PA I    T       VPHG HYH+IP                    + P +
Sbjct: 303  RHVESDGLVFDPAQITSRTARGVAVPHGDHYHFIPYSQMSELEEERIA-------RIIPLR 355

Query: 239  LSYSSTASDNNTQSVAKGSTSKPANKSENLQSLLKELYDSPSAQRYSE-SDGLVFDPAKI 297
                +    D+  +  +    T +P+   +   +L  +   S  +Q   +  +G VF+  K
Sbjct: 356  YRSNHWVPDSRLEQPSLQPTPEPSPGPQPAPNLKIDSNSSLVSQLVRKVGEGYVFEE-KG 414

Query: 298  ISRTPNGVAIPHGDHYHF---IPYSKLSALEEKIARRVPISGTGSTVSTNAKPNEVVXXX 354
            ISR             Y F    +P   +   LE K++++  +S T +    N   P +
Sbjct: 415  ISR-----------YVFAKDLPSETVKNLESKLSKQESVSHTLTAKKENVAPRDQEFYD 462

Query: 355  XXXXXXXXXXXXXXKELSSASDGYIFNPK--DIVEETATAYIVRHGDHFHYIPKSNQIGQP 412
                          E   +  +     K   + + + +T         D  ++       +
Sbjct: 463  KAYNLLTEAHKALFENKGRNSDFQALDKLLERLNDESTNKEKLVDDLLAFLAPITHPERL 522

Query: 413  TLPNNSLATPSPSLPINPGISHEKHEEDGYGFDANRIIAEDESGFIMSHGNHNHYYFFKKD 472
            PN+ +     +  I  ++ +      DGY FD + II+++   ++    H   H+H+    K
Sbjct: 523  GKPNSQIEYTEDEVRIAQ-LADKYTTSDGYIFDEHDIISDEGDAYVTLHMGHSHWIGKDS 581

Query: 473  LTEEQIKAAQKHLEEVKTSHNGLDSLSSHEQDYPGNAKEMKDLDKKIEEKIAGIMKQYGV 532
            L++++  AAQ + +E          D  + + +   G++           K E++I +   Y V
Sbjct: 582  LSDKEKVAAQAYTKEKGILPPSPD--ADVKANPTGDSAAAIYNRVKGEKRIPLVRLPYMV 639

Query: 533  KRESIVVNKEKNAIIYPHGDHHH 555
            +      V  +    +I PH DH+H
Sbjct: 640  EH---TVEVKNGNLIIPHKDHYH 659


  Score = 59.3 bits (142), Expect = 4e-07
  Identities = 76/340 (22%), Positives = 111/340 (32%), Gaps = 111/340 (32%)

Query: 239  LSYSSTASDNNTQSVAKGSTSKPANKSENLQSLLKELYDSPSAQRYSESDGLVFDPAKII 298
            L  + T  +NN S   G  +    K+ENL              +    S+ +G+  +   +I
Sbjct: 26   LYQARTVKENNRVSYIDGKQA--TQKTENL----------TPDEVSKREGINAEQI-VI 71

Query: 299  SRTPNGVAIPHGDHYHF----IPYSKLSALE------------EKIARRVP------ISG 336
              T  G    HGDHYH+    +PY + + E              E  I   V      + G
Sbjct: 72   KITDQGYVTSHGDHYHYYNGKVPYDAIFSEELLMKDPNYKLKDEDIVNEVKGGYVIKVDG 131

Query: 337  TGSTVSTNAKPNEVVXXXXXXXXXXXXXXXXXXXKELSS----------------ASDGYIF 379
              +A   + V                      +E  +                    DGYIF
Sbjct: 132  KYYVYLKDAAHADNVRTKEEINRQKQEHSQHREGGTPRNDGAVALARSQGRYTTDDGYIF 191

Query: 380  NPKDIVEETATAYIVRHGDHFHYIPK----------------------SNQIGQPTLPN 416
            N  DI+E+T  AYIV HGDH+HYIPK                       S   +   N
Sbjct: 192  NASDIIEDTGDAYIVPHGDHYHYIPKNELSASELAAAEAFLSGRGNLSNSRTYRRQNSDN 251

Query: 417  NSLATPSPSLPINPG-----------------------------------ISHEKHEEDGY 442
            S       PS+  NPG                                     +S    E DG
Sbjct: 252  TSRTNWVPSVS-NPGTTNTNTSNNSNTNGQASQSNEDVDSLLKQLYALPLSKRHVESDGL 310

Query: 443  GFDANRIIAEDESGFIMSHGNHNHYYFFKKDLTEEQIKAAQ 482
             FD  +I +    G + HG+H H+      ++E + + A+
Sbjct: 311  VFDPAQITSRTARGVAVPHGDHYHFIPYSQMSELEEERIAR 350
```

```
 Score = 53.1 bits (126), Expect = 3e-05
 Identities = 35/155 (22%), Positives = 56/155 (35%), Gaps = 20/155 (12%)

Query: 165 EKVNSNVAVARSQGRYTTNDGYVFNPADIIEDTGNAYIVPHGGHYHYIPXXXXXXXXXXX 224
            E    V +A+    +YTT+DGY+F+  DII D G+AY+  H GH H+I
Sbjct: 530 EYTEDEVRIAQLADKYTTSDGYIFDEHDIISDEGDAYVTHMGHSHWIGKDSLSDKEKVA 589

Query: 225 XXXXXXXXXNMQPSQLSYSSTASDNNTQSVAKGSTSKPANKSENLQSLLKELYDSPSAQRY 284
                     + P       A+      + A                        +Y+    ++
Sbjct: 590 AQAYTKEKGILPPSPDADVKANPTGDSAAA--------------------IYNRVKGEKR 629

Query: 285 SESDGLVFDPAKIISRTPNGVAIPHGDHYHFIPYS 319
                   L +     +     + IPH DHYH I ++
Sbjct: 630 IPLVRLPYMVEHTVEVKNGNLIIPHKDHYHNIKFA 664



 Score = 43.5 bits (101), Expect = 0.021
 Identities = 52/271 (19%), Positives = 94/271 (34%), Gaps = 93/271 (34%)

Query: 384 IVEETATAYIVRHGDHFHY----IPKSNQIGQPTL---PNNSLATPSPSLPINPGI---- 432
            +++ T  Y+ HGDH+HY    +P      + L  PN L        + G
Sbjct: 70  VIKITDQGYVTSHGDHYHYYNGKVPYDAIFSEELLMKDPNYKLKDEDIVNEVKGGYVIKV 129

Query: 433 --------------------------SHEKHEEDG--------------------Y 442
                                      H +H E G                    Y
Sbjct: 130 DGKYYVYLKDAAHADNVRTKEEINRQKQEHSQHREGGTPRNDGAVALARSQGRYTTDDGY 189

Query: 443 GFDANRIIAEDESGFIMSHGNHNHYFFKKDLTEEQIKAAQKHL---------------E 486
            F+A+ II +     +I+ HG+H HY  K +L+  ++ AA+  L
Sbjct: 190 IFNASDIIEDTGDAYIVPHGDHYHYIPKNELSASELAAAEAFLSGRGNLSNSRTYRRQNS 249

Query: 487 EVKTSHNGLDSLS---------SHEQDYPGNAKEM-KDLDKKIEEKIAGIMKQYGVKRES 536
            + + N + S+S         S+ + G A + +D+D +++ A + V+ +
Sbjct: 250 DNTSRTNWVPSVSNPGTTNTNTSNNSNTNGQASQSNEDVDSLLKQLYALPLSKRHVESDG 309

Query: 537 IV------VNKEKNAIIYPHGDHHHADPIDE 561
            +V      ++     + PHGDH+H P +
Sbjct: 310 LVFDPAQITSRTARGVAVPHGDHYHFIPYSQ 340
```

tr Q97QM8 **Conserved domain protein [SP1175] [Streptococcus pneumoniae]** 802 AA

align

```
 Score =  424 bits (1089), Expect = e-117
 Identities = 209/352 (59%), Positives = 259/352 (73%), Gaps = 20/352 (5%)

Query: 18  LSLCAYALNQHRSQENKDNNRVSYVDGSQSSQKSENLTPDQVSQKEGIQAEQIVIKITDQ 77
            LS+C+Y L ++++  K+NNRVSY+DG Q++QK+ENLTPD+VS++EGI AEQIVIKITDQ
Sbjct: 3   LSVCSYELGLYQARTVKENNRVSYIDGKQATQKTENLTPDEVSKREGINAEQIVIKITDQ 62

Query: 78  GYVTSHGDHYHYYNGKVPYDALFSEELLMKDPNYQLKDADIVNEVKGGYIIKVDGKYYVY 137
            GYVTSHGDHYHYYNGKVPYDA+ SEELLMKDPNY+LKD DIVNEVKGGY+IKVDGKYYVY
Sbjct: 63  GYVTSHGDHYHYYNGKVPYDAIISEELLMKDPNYKLKDEDIVNEVKGGYVIKVDGKYYVY 122
```

```
Query: 138  LKDAAHADNVRTKDEINRQKQEHVKDNE----KVNSNVAVARSQGRYTTNDGYVFNPADI 193
            LKDAAHADNVRTK+EINRQKQEH + E   + +  VA+ARSQGRYTT+DGY+FN +DI
Sbjct: 123  LKDAAHADNVRTKEEINRQKQEHSQHREGGTPRNDGAVALARSQGRYTTDDGYIFNASDI 182

Query: 194  IEDTGNAYIVPHGGHYHYIPXXXXXXXXXXXXXXXXXXXXXNMQPSQLSYSSTASDN----- 248
            IEDTG+AYIVPHG HYHYIP                      +  +Y    SDN
Sbjct: 183  IEDTGDAYIVPHGDHYHYIPKNELSASELAAAEAFLSGRGNLSNSRTYRRQNSDNTSRTN 242

Query: 249  ----------NTQSVAKGSTSKPANKSENLQSLLKELYDSPSAQRYSESDGLVFDPAKI 297
                      NT +    +T+ A++S ++ SLLK+LY  P +QR+ ESDGLVFDPA+I
Sbjct: 243  WVPSVSNPGTTNTNTSNNSNTNSQASQSNDIDSLLKQLYKLPLSQRHVESDGLVFDPAQI 302

Query: 298  ISRTPNGVAIPHGDHYHFIPYSKLSALEEKIARRVPISGTGSTVSTNAKPNE 349
            SRT  GVA+PHGDHYHFIPYS++S LEE+IAR +P+    +   +++P +
Sbjct: 303  TSRTARGVAVPHGDHYHFIPYSQMSELEERIARIIPLRYRSNHWVPDSRPEQ 354
```

```
 Score = 79.0 bits (193), Expect = 5e-13
 Identities = 65/282 (23%), Positives = 108/282 (38%), Gaps = 57/282 (20%)
```

```
Query: 280  SAQRYSESDGLVFDPAKIISRTPNGVAIPHGDHYHFIPYSKLSALEEKIARRVPISGTGS 339
            S RY+  DG +F+ + II T +     +PHGDHYH+IP ++LSA  E  A    +SG G+
Sbjct: 165  SQGRYTTDDGYIFNASDIIEDTGDAYIVPHGDHYHYIPKNELSA-SELAAAEAFLSGRGN 223

Query: 340  ---------------TVSTN-----AKPNEVVXXXXXXXXXXXXXXXXXXXKELSSA------- 373
                           T  TN     + P                        ++ S
Sbjct: 224  LSNSRTYRRQNSDNTSRTNWVPSVSNPGTTNTNTSNNSNTNSQASQSNDIDSLLKQLYKL 283

Query: 374  --------SDGYIFNPKDIVEETATAYIVRHGDHFHYIPKSNQI---------------G 410
                    SDG +F+P  I    TA   V HGDH+H+IP S                 G
Sbjct: 284  PLSQRHVESDGLVFDPAQITSRTARGVAVPHGDHYHFIPYSQMSELEERIARIIPLRYRS 343

Query: 411  QPTLPNNSLATPSPSLPINPGISHEKHEE---DGYGFDANRIIAEDESGFIMSHGNHNHY 467
            +P++     PSP  P  +         D       ++++ +   G++      + Y
Sbjct: 344  NHWVPDSRPEQPSPQPTPEPSPGPQPAPNLKIDSNSSLVSQLVRKVGEGYVFEEKGISRY 403

Query: 468  FFKKDLTEEQIKAAQKHLEEVKTSHNGL----DSLSSHEQDY 505
            F KDL E +K + L + ++  + L    +++  +Q++
Sbjct: 404  VFAKDLPSETVKNLESKLSKQESVSHTLTAKKENVAPRDQEF 445
```

```
 Score = 77.8 bits (190), Expect = 1e-12
 Identities = 103/504 (20%), Positives = 189/504 (37%), Gaps = 60/504 (11%)
```

```
Query: 71   VIKITDQGYVTSHGDHYHYYNGKVPYDALFSEEL-----------LMKDPNYQLKDADI 118
            +I+ T  Y+  HGDHYHY   +P + L + EL         L     Y+ +++D
Sbjct: 182  IIEDTGDAYIVPHGDHYHY----IPKNELSASELAAAEAFLSGRGNLSNSRTYRRQNSD- 236

Query: 119  VNEVKGGYIIKVDGKYYVYLKDAAHADNVRTKDEINRQKQ-EHVKDNEKVNSNVAVARSQ 177
            N  +  ++ V         N T + N Q    D + +
Sbjct: 237  -NTSRTNWVPSVSNP---------GTTNTNTSNNSNTNSQASQSNDIDSLLKQLYKLPLS 286

Query: 178  GRYTTNDGYVFNPADIIEDTGNAYIVPHGGHYHYIPXXXXXXXXXXXXXXXXXXXXXNMQPS 237
            R+  +DG VF+PA I   T      VPHG HYH+IP                     + P
Sbjct: 287  QRHVESDGLVFDPAQITSRTARGVAVPHGDHYHFIPYSQMSELEERIA-------RIIPL 339

Query: 238  QLSYSSTASDNNTQSVAKGSTSKPANKSENLQSLLKELYDSPSAQRYSE-SDGLVFDPAK 296
```

```
                +    +     D+  +   +    T +P+   +   +L +   S  +Q   +  +G VF+  K
Sbjct: 340 RYRSNHWVPDSRPEQPSPQPTPEPSPGPQPAPNLKIDSNSSLVSQLVRKVGEGYVFEE-K 398

Query: 297 IISRTPNGVAIPHGDHYHF---IPYSKLSALEEKIARRVPISGTGSTVSTNAKPNEVVXX 353
               ISR          Y F  +P  +  LE K++++ +S T +   N  P +
Sbjct: 399 GISR-----------YVFAKDLPSETVKNLESKLSKQESVSHTLTAKKENVAPRDQEFY 446

Query: 354 XXXXXXXXXXXXXXXKELSSASDGYIFNPK--DIVEETATAYIVRHGDHFHYIPKSNQIGQ 411
                           E   +  +    K  + + +T     D  ++          +
Sbjct: 447 DKAYNLLTEAHKALFENKGRNSDFQALDKLLERLNDESTNKEKLVDDLLAFLAPITHPER 506

Query: 412 PTLPNNSLATPSPSLPINPGISHEKHEEDGYGFDANRIIAEDESGFIMSHGNHNHYFFKK 471
               PN+ +    + I  ++ +     DGY FD + II+++  ++  H  H+H+   K
Sbjct: 507 LGKPNSQIEYTEDEVRIAQ-LADKYTTSDGYIFDEHDIISDEGDAYVTPHMGHSHWIGKD 565

Query: 472 DLTEEQIKAAQKHLEEVKTSHNGLDSLSSHEQDYPGNAKEMKDLDKKIEEKIAGIMKQYG 531
               L++++  AAQ + +E         D  +  +  G++        K E++I  +    Y
Sbjct: 566 SLSDKEKVAAQAYTKEKGILPPSPD--ADVKANPTGDSAAAIYNRVKGEKRIPLVRLPYM 623

Query: 532 VKRESIVVNKEKNAIIYPHGDHHH 555
               V+      V +   +I PH DH+H
Sbjct: 624 VEH---TVEVKNGNLIIPHKDHYH 644


  Score = 60.1 bits (144), Expect = 2e-07
  Identities = 73/372 (19%), Positives = 131/372 (34%), Gaps = 107/372 (28%)

Query: 239 LSYSSTASDNNTQSVAKGSTSKPANKSENLQSLLKELYDSPSAQRYSESDGLVFDPAKII 298
               L + T +NN  S   G   +    K+ENL             +    S+ +G+  +    +I
Sbjct: 12  LYQARTVKENNRVSYIDGKQA--TQKTENL-----------TPDEVSKREGINAEQI-VI 57

Query: 299 SRTPNGVAIPHGDHYHF----IPYSKLSALEEKIARRVPISGTGSTVSTNAKPNEVVXXX 354
               T G    HGDHYH+    +PY + +
Sbjct: 58  KITDQGYVTSHGDHYHYYNGKVPYDAIIS----------------------------- 86

Query: 355 XXXXXXXXXXXXXXXKELSSASDGYIFNPKDIVEETATAYIVR------------HGDHFH 401
                             +EL    Y   +DIV E    Y+++             H D+
Sbjct: 87  ------------EELLMKDPNYKLKDEDIVNEVKGGYVIKVDGKYYVYLKDAAHADNVR 133

Query: 402 YIPKSNQIGQPTLPNNSLATPSPSLPINPGISHEKHE-EDGYGFDANRIIAEDESGFIMS 460
               + N+ Q   +    TP   +    S ++ +DGY F+A+ II +     +I+
Sbjct: 134 TKEEINRQKQEHSQHREGGTPRNDGAVALARSQGRYTTDDGYIFNASDIIEDTGDAYIVP 193

Query: 461 HGNHNHYFFKKDLTEEQIKAAQKHL---------------EEVKTSHNGLDSLS----- 499
               HG+H HY  K +L+ ++ AA+ L               +  +  N + S+S
Sbjct: 194 HGDHYHYIPKNELSASELAAAEAFLSGRGNLSNSRTYRRQNSDNTSRTNWVPSVSNPGTT 253

Query: 500 ----SHEQDYPGNAKEMKDLDKKIEEKIAGIMKQYGVKRESIV------VNKEKNAIIYP 549
               S+  +    A +  D+D +++       + Q V+  + +V          ++   +  P
Sbjct: 254 NTNTSNNSNTNSQASQSNDIDSLLKQLYKLPLSQRHVESDGLVFDPAQITSRTARGVAVP 313

Query: 550 HGDHHHADPIDE 561
               HGDH+H  P  +
Sbjct: 314 HGDHYHFIPYSQ 325


  Score = 57.0 bits (136), Expect = 2e-06
  Identities = 36/155 (23%), Positives = 57/155 (36%), Gaps = 20/155 (12%)
```

```
Query: 165 EKVNSNVAVARSQGRYTTNDGYVFNPADIIEDTGNAYIVPHGGHYHYIPXXXXXXXXXXXX 224
            E    V +A+   +YTT+DGY+F+  DII D G+AY+ PH GH H+I
Sbjct: 515 EYTEDEVRIAQLADKYTTSDGYIFDEHDIISDEGDAYVTPHMGHSHWIGKDSLSDKEKVA 574

Query: 225 XXXXXXXXXNMQPSQLSYSSTASDNNTQSVAKGSTSKPANKSENLQSLLKELYDSPSAQRY 284
                    + P      A+      + A                    +Y+    ++
Sbjct: 575 AQAYTKEKGILPPSPDADVKANPTGDSAAA-------------------IYNRVKGEKR 614

Query: 285 SESDGLVFDPAKIISRTPNGVAIPHGDHYHFIPYS 319
               L +    +     + IPH DHYH I ++
Sbjct: 615 IPLVRLPYMVEHTVEVKNGNLIIPHKDHYHNIKFA 649
```

tr Q9ANY2 **Pneumococcal histidine triad protein D precursor (Fragment)** 839 AA
          (Hypothetical protein SP1003) (Fragment) [phtD]
          [Streptococcus pneumoniae]   align

```
 Score = 423 bits (1088), Expect = e-117
 Identities = 215/359 (59%), Positives = 268/359 (73%), Gaps = 14/359 (3%)

Query: 1   MKFSKKYIAAGSAVIVSLSLCAYALNQHRS-QENKDNNRVSYVDGSQSSQKSENLTPDQV 59
           MK +KKY+A GS  +++LS+C+Y L +H++ Q  K++NRVSY+DG Q+ QK+ENLTPD+V
Sbjct: 1   MKINKKYLA-GSVAVLALSVCSYELGRHQAGQVKKESNRVSYIDGDQAGQKAENLTPDEV 59

Query: 60  SQKEGIQAEQIVIKITDQGYVTSHGDHYHYYNGKVPYDALFSEELLMKDPNYQLKDADIV 119
           S++EGI AEQIVIKITDQGYVTSHGDHYHYYNGKVPYDA+ SEELLMKDPNYQLKD+DIV
Sbjct: 60  SKREGINAEQIVIKITDQGYVTSHGDHYHYYNGKVPYDAIISEELLMKDPNYQLKDSDIV 119

Query: 120 NEVKGGYIIKVDGKYYVYLKDAAHADNVRTKDEINRQKQEHVKDNEKVNSN---VAVARS 176
           NE+KGGY+IKVDGKYYVYLKDAAHADN+RTK+EI RQKQEH    N   SN   V AR+
Sbjct: 120 NEIKGGYVIKVDGKYYVYLKDAAHADNIRTKEEIKRQKQEH-SHNHGGGSNDQAVVAARA 178

Query: 177 QGRYTTNDGYVFNPADIIEDTGNAYIVPHGGHYHYIPXXXXXXXXXXXXXXXXXXXXNMQP 236
           QGRYTT+DGY+FN +DIIEDTG+AYIVPHG HYHYIP                    Q
Sbjct: 179 QGRYTTDDGYIFNASDIIEDTGDAYIVPHGDHYHYIP--KNELSASELAAAEAYWNGKQG 236

Query: 237 SQLSYSSTASDNNTQ---SVAKGSTSKPA---NKSENLQSLLKELYDSPSAQRYSESDGL 290
           S+ S SS+ + N Q   S    T  P    N+ EN+ SLL+ELY  P ++R+ ESDGL
Sbjct: 237 SRPSSSSSYNANPAQPRLSENHNLTVTPTYHQNQGENISSLLRELYAKPLSERHVESDGL 296

Query: 291 VFDPAKIISRTPNGVAIPHGDHYHFIPYSKLSALEEKIARRVPISGTGSTVSTNAKPNE 349
           +FDPA+I SRT  GVA+PHG+HYHFIPY ++S LE++IAR +P+  +     +++P +
Sbjct: 297 IFDPAQITSRTARGVAVPHGNHYHFIPYEQMSELEKRIARIIPLRYRSNHWVPDSRPEQ 355
```

```
 Score = 77.8 bits (190), Expect = 1e-12
 Identities = 106/501 (21%), Positives = 185/501 (36%), Gaps = 62/501 (12%)

Query: 71  VIKITDQGYVTSHGDHYHYYNGKVPYDALFSEELLMKDPNYQLKDADIVNEVKGGYIIKV 130
           +I+ T  Y+ HGDHYHY     +P + L + EL  + +  K      +
Sbjct: 195 IIEDTGDAYIVPHGDHYHY----IPKNELSASELAAAEAYWNGKQGSRPSSSSS------ 244

Query: 131 DGKYYVYLKDAAHADNVRTKDEINRQKQEHVKDNEKVNSNVAVARSQG---RYTTNDGYV 187
             Y  + A    +     + H   E ++S +   ++ R+ +DG +
Sbjct: 245 ------YNANPAQP-RLSENHNLTVTPTYHQNQGENISSLLRELYAKPLSERHVESDGLI 297
```

```
Query:  188  FNPADIIEDTGNAYIVPHGGHYHYIPXXXXXXXXXXXXXXXXXXXXX-------NMQPSQLS  240
             F+PA I    T     VPHG HYH+IP                           + +P Q S
Sbjct:  298  FDPAQITSRTARGVAVPHGNHYHFIPYEQMSELEKRIARIIPLRYRSNHWVPDSRPEQPS  357

Query:  241  YSSTASDNNTQSVAKGSTSKPANKSENLQSLLKELYDSPSAQRYSESDGLV-FDPAKIIS  299
             ST   + +   A     P+N  +  + L+KE          E +G+  + PAK +S
Sbjct:  358  PQSTPEPSPSPQPAPNPQPAPSNPID--EKLVKEAVRKVGDGYVFEENGVSRYIPAKDLS  415

Query:  300  -RTPNGVAIPHGDHYHFIPYSKLSALEEKIARRVPISGTGSTVSTNAKPNEVVXXXXXXX  358
              T  G+           SKL A +E ++ ++    T   S     N+
Sbjct:  416  AETAAGID-----------SKL-AKQESLSHKLGAKKTDLPSSDREFYNKAYDLLARIH  462

Query:  359  XXXXXXXXXXKELSSASDGYIFNPKDIVEETATAYIVRHGDHFHYIPKSNQIGQPTLPNNS  418
                      +    A D +   KD+ +           D  ++      +    PN
Sbjct:  463  QDLLDNKGRQVDFEALDNLLERLKDVPSDKVKLV----DDILAFLAPIRHPERLGKPNAQ  518

Query:  419  LATPSPSLPINPGISHEKHEEDGYGFDANRIIAEDESGFIMSHGNHNHYYFFKKDLTEEQI  478
             +      + +  ++ +    EDGY FD  I +++   ++  H H+H+  K  L+E +
Sbjct:  519  ITYTDDEIQVAK-LAGKYTTEDGYIFDPRDITSDEGDAYVTPHMTHSHWIKKDSLSEAER  577

Query:  479  KAAQKHLEEVKTSHNGLDSLSSHEQDYPGN--AKEMKDLDKKIE--EKIAGIMKQYGVKR  534
             AAQ + +E     GL  S+  QD GN  AK  + +  +++  +K+      Y ++
Sbjct:  578  AAAQAYAKE-----KGLTPPSTDHQD-SGNTEAKGAEAIYNRVKAAKKVPLDRMPYNLQ-  630

Query:  535  ESIVVNKEKNAIIYPHGDHHH  555
              V  +   ++I PH DH+H
Sbjct:  631  --YTVEVKNGSLIIPHYDHYH  649
```

 Score = 77.0 bits (188), Expect = 2e-12
 Identities = 72/276 (26%), Positives = 109/276 (39%), Gaps = 54/276 (19%)

```
Query:  283  RYSESDGLVFDPAKIISRTPNGVAIPHGDHYHFIPYSKLSALEEKIARRVPISGTGSTVS  342
             RY+  DG +F+ + II  T +    +PHGDHYH+IP ++LSA E   A      GS  S
Sbjct:  181  RYTTDDGYIFNASDIIEDTGDAYIVPHGDHYHYIPKNELSASELAAAEAYWNGKQGSRPS  240

Query:  343  T----NAKP---------NEVVXXXXXXXXXXXXXXXXXKELSSA--------SDGYIFNP  381
             +    NA P         N V                 +EL +        SDG IF+P
Sbjct:  241  SSSSYNANPAQPRLSENHNLTVTPTYHQNQGENISSLLRELYAKPLSERHVESDGLIFDP  300

Query:  382  KDIVEETATAYIVRHGDHFHYIP------------------KSN------QIGQPTLPN  416
             I   TA   V HG+H+H+IP                  +SN      + QP+  +
Sbjct:  301  AQITSRTARGVAVPHGNHYHFIPYEQMSELEKRIARIIPLRYRSNHWVPDSRPEQPSPQS  360

Query:  417  NSLATPSPSLPINPGISHEKHEEDGYGFDANRIIAEDESGFIMSHGNHNHYYFFKKDLTEE  476
             +PSP   NP +      ++    +A R + +     G++        + Y  KDL+ E
Sbjct:  361  TPEPSPSPQPAPNPQPAPSNPIDEKLVKEAVRKVGD---GYVFEENGVSRYIPAKDLSAE  417

Query:  477  QIKAAQKHL-EEVKTSHNGLDSLSSHEQDYPGNAKE  511
                L ++    SH      L + + D P + +E
Sbjct:  418  TAAGIDSKLAKQESLSH----KLGAKKTDLPSSDRE  449
```

 Score = 67.4 bits (163), Expect = 1e-09
 Identities = 49/190 (25%), Positives = 79/190 (40%), Gaps = 30/190 (15%)

```
Query:  138  LKDAAHADNVRTKDEIN------RQKQEHVKDNEKV---NSNVAVARSQGRYTTNDGYVF  188
```

```
             LKD    +D V+  D+I       R  +   K N ++     +  + VA+  G+YTT DGY+F
Sbjct: 485 LKDVP-SDKVKLVDDILAFLAPIRHPERLGKPNAQITYTDDEIQVAKLAGKYTTEDGYIF 543

Query: 189 NPADIIEDTGNAYIVPHGGHYHYIPXXXXXXXXXXXXXXXXXXXXXNMQPSQLSYSSTASDN 248
             +P DI  D G+AY+ PH  H H+I                       + P      ST   +
Sbjct: 544 DPRDITSDEGDAYVTPHMTHSHWIKKDSLSEAERAAAQAYAKEKGLTP-----PSTDHQD 598

Query: 249 NTQSVAKGSTSKPANKSENLQSLLKELYDSPSAQRYSESDGLVFDPAKIISRTPNGVAIP 308
             +  + AKG+     + +Y+     A +    D + ++    +       + IP
Sbjct: 599 SGNTEAKGA--------------EAIYNRVKAAKKVPLDRMPYNLQYTVEVKNGSLIIP 643

Query: 309 HGDHYHFIPY 318
             H DHYH I +
Sbjct: 644 HYDHYHNIKF 653


  Score = 48.5 bits (114), Expect = 7e-04
  Identities = 51/288 (17%), Positives = 104/288 (35%), Gaps = 55/288 (19%)

Query: 297 IISRTPNGVAIPHGDHYHF----IPYSKLSALEEKIARRVPISGTGSTVSTNAKPNEVVX 352
             +I  T  G   HGDHYH+    +PY + + E++ + P        +   K +++V
Sbjct: 71  VIKITDQGYVTSHGDHYHYYNGKVPYDAI--ISEELLMKDP--------NYQLKDSDIVN 120

Query: 353 XXXXXXXXXXXXXXXXKELSSASDGYIFNPKDIVEETATAYIVRHGDHFHYIPKSNQIGQP 412
                            L  A+      K+ ++     + HG        ++Q
Sbjct: 121 EIKGGYVIKVDGKYYVYLKDAAHADNIRTKEEIKRQKQEHSHNHGGG-----SNDQAVVA 175

Query: 413 TLPNNSLATPSPSLPINPGISHEKHEEDGYGFDANRIIAEDESGFIMSHGNHNHYFFKKD 472
                 T                +DGY F+A+ II +    +I+ HG+H HY  K +
Sbjct: 176 ARAQGRYTT---------------DDGYIFNASDIIEDTGDAYIVPHGDHYHYIPKNE 218

Query: 473 LTEEQIKAAQKHLEEVKTSHNGLDSLSSHEQDYPG-------------NAKEMKDLDKKI 519
             L+ ++ AA+ +    + S    S +   P               + + +++   +
Sbjct: 219 LSASELAAAEAYWNGKQGSRPSSSSSYNANPAQPRLSENHNLTVTPTYHQNQGENISSLL 278

Query: 520 EEKIAGIMKQYGVKRESI------VVNKEKNAIIYPHGDHHHADPIDE 561
             E  A + + V+ + +        + ++   + PHG+H+H  P ++
Sbjct: 279 RELYAKPLSERHVESDGLIFDPAQITSRTARGVAVPHGNHYHFIPYEQ 326


  Score = 42.0 bits (97), Expect = 0.061
  Identities = 44/185 (23%), Positives = 64/185 (33%), Gaps = 36/185 (19%)

Query: 372 SASDGYIFNPKDIVEETATAYIVRHGDHFHYIPKSN------QIGQPTLPNNSLATPSPS 425
             +  DGYIF+P+DI  +   AY+  H  H H+I K +         Q       L   PS
Sbjct: 536 TTEDGYIFDPRDITSDEGDAYVTPHMTHSHWIKKDSLSEAERAAAQAYAKEKGLTPPSTD 595

Query: 426 LPINPGISHEKHEE--------------DGYGFDANRIIAEDESGFIMSHGNHNH----- 466
             + G + K E             D  ++    +         I+ H +H H
Sbjct: 596 HQ-DSGNTEAKGAEAIYNRVKAAKKVPLDRMPYNLQYTVEVKNGSLIIPHYDHYHNIKFE 654

Query: 467 ------YFFKKDLTEEQIKAAQK----HLEEVKTSHNGLDSLSSHEQDYPGNAKEMKDLD 516
                   Y  K  T E +A K      H E   SNG  + SH +    + D  D
Sbjct: 655 WFDEGLYEAPKGYTLEDLLATVKYYVEHPNERPHSDNGFGNASDHVRKNKVDQDSKPDED 714

Query: 517 KKIEE 521
             K+ +E
Sbjct: 715 KEHDE 719
```

```
Score = 39.7 bits (91), Expect = 0.30
Identities = 34/138 (24%), Positives = 51/138 (36%), Gaps = 19/138 (13%)

Query: 274  ELYDSPSAQRYSESDGLVFDPAKIISRTPNGVAIPHGDHYHFIPYSKLSALEEKIARRV- 332
            E+  +  A  Y+  DG +FDP   I S   +    PH H H+I   LS E   A+
Sbjct: 525  EIQVAKLAGKYTTEDGYIFDPRDITSDEGDAYVTPHMTHSHWIKKDSLSEAERAAAQAYA 584

Query: 333  -------PISGTGSTVSTNAKPNEVVXXXXXXXXXXXXXXXXXXXKELSSASDGYIFNPKDIV 385
                   P +    + +T AK  E +                      D  +N +  V
Sbjct: 585  KEKGLTPPSTDHQDSGNTEAKGAEAIYNRVKAAK----------KVPLDRMPYNLQYTV 633

Query: 386  EETATAYIVRHGDHFHYI 403
            E    + I+ H DH+H I
Sbjct: 634  EVKNGSLIIPHYDHYHNI 651
```

tr Q97QM9 **Conserved domain protein [SP1174] [Streptococcus pneumoniae]** 819 AA

align

```
Score =  423 bits (1087), Expect = e-116
Identities = 211/357 (59%), Positives = 271/357 (75%), Gaps = 11/357 (3%)

Query: 1    MKFSKKYIAAGSAVIVSLSLCAYALNQHRS-QENKDNNRVSYVDGSQSSQKSENLTPDQV 59
            MK +KKY+A GS  +++LS+C+Y L ++++ Q+ K++NRV+Y+DG Q+ QK+ENLTPD+V
Sbjct: 1    MKINKKYLA-GSVAVLALSVCSYELGRYQAGQDKKESNRVAYIDGDQAGQKAENLTPDEV 59

Query: 60   SQKEGIQAEQIVIKITDQGYVTSHGDHYHYYNGKVPYDALFSEELLMKDPNYQLKDADIV 119
            S++EGI AEQIVIKITDQGYVTSHGDHYHYYNGKVPYDA+ SEELLMKDPNYQLKD+DIV
Sbjct: 60   SKREGINAEQIVIKITDQGYVTSHGDHYHYYNGKVPYDAIISEELLMKDPNYQLKDSDIV 119

Query: 120  NEVKGGYIIKVDGKYYVYLKDAAHADNVRTKDEINRQKQEHVKD-NEKVNSNVAVARSQG 178
            NE+KGGY+IKV+GKYYVYLKDAAHADN+RTK+EI RQKQE   + N + + ++ VA AR+QG
Sbjct: 120  NEIKGGYVIKVNGKYYVYLKDAAHADNIRTKEEIKRQKQERSHNHNSRADNAVAAARAQG 179

Query: 179  RYTTNDGYVFNPADIIEDTGNAYIVPHGGHYHYIPXXXXXXXXXXXXXXXXXXXXXNMQPSQ 238
            RYTT+DGY+FN +DIIEDTG+AYIVPHG HYHYIP                      Q S+
Sbjct: 180  RYTTDDGYIFNASDIIEDTGDAYIVPHGDHYHYIP--KNELSASELAAAEAYWNGKQGSR 237

Query: 239  LSYSSTASDNNTQ---SVAKGSTSKPA---NKSENLQSLLKELYDSPSAQRYSESDGLVF 292
            S SS+ + N Q   S        T  P   N+ EN+ SLL+ELY  P ++R+ ESDGL+F
Sbjct: 238  PSSSSYNANPAQPRLSENHNLTVTPTYHQNQGENISSLLRELYAKPLSERHVESDGLIF 297

Query: 293  DPAKIISRTPNGVAIPHGDHYHFIPYSKLSALEEKIARRVPISGTGSTVSTNAKPNE 349
            DPA+I SRT GVA+PHG+HYHFIPY ++S LE++IAR +P+     +    +++P E
Sbjct: 298  DPAQITSRTARGVAVPHGNHYHFIPYEQMSELEKRIARIIPLRYRSNHWVPDSRPEE 354
```

```
Score = 82.0 bits (201), Expect = 5e-14
Identities = 93/391 (23%), Positives = 135/391 (33%), Gaps = 91/391 (23%)

Query: 193  IIEDTGNAYIVPHGGHYHYI------------------PXXXXXXXXXXXXXXXXXXXXXNM 234
            +I+ T  Y+  HG HYHY                   P                     +
Sbjct: 71   VIKITDQGYVTSHGDHYHYYNGKVPYDAIISEELLMKDPNYQLKDSDIVNEIKGGYVIKV 130
```

```
Query: 235 QPSQLSYSSTASDNNTQSVAKGSTSKPANKSENLQSLLKELYDSPSAQ-RYSESDGLVFD 293
                  Y  A+  +      +      +   +S N  S         +  AQ RY+   DG +F+
Sbjct: 131 NGKYYVYLKDAAHADNIRTKEEIKRQKQERSHNHNSRADNAVAAARAQGRYTTDDGYIFN 190

Query: 294 PAKIISRTPNGVAIPHGDHYHFIPYSKLSALEEKIARRVPISGTGSTVST----NAKP-- 347
             +  II   T +    +PHGDHYH+IP ++LSA E    A      GS   S+    NA P
Sbjct: 191 ASDIIEDTGDAYIVPHGDHYHYIPKNELSASELAAAEAYWNGKQGSRPSSSSSYNANPAQ 250

Query: 348 -------NEVVXXXXXXXXXXXXXXXXXXXKELSSA--------SDGYIFNPKDIVEETATAY 392
                   N  V                    +EL +      SDG IF+P  I    TA
Sbjct: 251 PRLSENHNLTVTPTYHQNQGENISSLLRELYAKPLSERHVESDGLIFDPAQITSRTARGV 310

Query: 393 IVRHGDHFHYIP------------------KSNQ---------IGQPTLPNNSLATPS 423
              V HG+H+H+IP                    +SN           QPT    +    P+
Sbjct: 311 AVPHGNHYHFIPYEQMSELEKRIARIIPLRYRSNHWVPDSRPEEPSPQPTPEPSPSPQPA 370

Query: 424 PSLPINPGISHE--KHEEDGYGFDANRIIAEDESGFIMSHGNHNHYFFKKDLTEEQIKAA 481
            PS PI+  +  E  +    DGY F+ N +         + Y   KDL+ E
Sbjct: 371 PSNPIDEKLVKEAVRKVGDGYVFEENGV--------------SRYIPAKDLSAETAAGI 415

Query: 482 QKHL-EEVKTSHNGLDSLSSHEQDYPGNAKE 511
              L ++    SH       L + + D P + +E
Sbjct: 416 DSKLAKQESLSH----KLGTKKTDLPSSDRE 442


  Score = 76.3 bits (186), Expect = 3e-12
  Identities = 116/555 (20%), Positives = 213/555 (37%), Gaps = 95/555 (17%)

Query: 29  RSQENKDNNRVSYVDGSQSSQKSENLTPDQVSQKEG-IQAEQIVIKITDQGYVTSHGDHY 87
            R ++ + +N S  D + ++ +++       +  +G I      +I+ T   Y+   HGDHY
Sbjct: 155 RQKQERSHNHNSRADNAVAAARAQG----RYTTDDGYIFNASDIIEDTGDAYIVPHGDHY 210

Query: 88  HYYNGKVPYDALFSEELLMKDPNYQLKDADIVNEVKGGYIIKVDGKYYVYLKDAAHADNV 147
            HY     +P + L + EL      + K          +       Y   + A     +
Sbjct: 211 HY----IPKNELSASELAAAEAYWNGKQGSRPSSSS------------YNANPAQP-RL 253

Query: 148 RTKDEINRQKQEHVKDNEKVNSNVAVARSQG---RYTTNDGYVFNPADIIEDTGNAYIVP 204
                 +    H    E ++S +      ++     R+  +DG +F+PA I   T      VP
Sbjct: 254 SENHNLTVTPTYHQNQGENISSLLRELYAKPLSERHVESDGLIFDPAQITSRTARGVAVP 313

Query: 205 HGGHYHYIPXXXXXXXXXXXXXXXXXXXXXXXNMQPSQLSYSSTASDNNTQSVAKGSTS-KPAN 263
            HG HYH+IP                           + +      S   + Q    + S S +PA
Sbjct: 314 HGNHYHFIPYEQMSELEKRIARIIPL--RYRSNHWVPDSRPEEPSPQPTPEPSPSPQPAP 371

Query: 264 KSENLQSLLKELYDSPSAQRYSESDGLVFDPAKIISRTPNGVAIPHGDHYHFIPYSKLSA 323
             +   + L+KE         DG VF+        NGV+        +IP   LSA
Sbjct: 372 SNPIDEKLVKEAVRKVG-------DGYVFEE--------NGVS-------RYIPAKDLSA 409

Query: 324 -----LEEKIARRVPISGTGSTVSTNAKP------NEVVXXXXXXXXXXXXXXXXXXXKELSS 372
                 ++ K+A++  +S    T  T+           N+                     +
Sbjct: 410 ETAAGIDSKLAKQESLSHKLGTKKTDLPSSDREFYNKAYDLLARIHQDLLDNKGRQVDFE 469

Query: 373 ASDGYIFNPKDI------VEETATAYI--VRHGDHFHYIPKSNQIGQPTLPNNSLATPSP 424
            A D  + KD+        + E   A++ +RH +            ++G+P   N  +
Sbjct: 470 ALDNLLERLKDVSSDKVKLVEDILAFLAPIRHPE--------RLGKP---NAQITYTDD 517

Query: 425 SLPINPGISHEKHEEDGYGFDANRIIAEDESGFIMSHGNHNHYFFKKDLTEEQIKAAQKH 484
```

```
                  + +    ++ +   EDGY FD   I +++    ++  H  H+H+  K  L+E +  AAQ +
Sbjct: 518 EIQVAK-LAGKYTTEDGYIFDPRDITSDEGDAYVTPHMTHSHWIKKDSLSEAERAAAQAY 576

Query: 485 LEEVKTSHNGLDSLSSHEQDYPGN--AKEMKDLDKKIE--EKIAGIMKQYGVKRESIVVN 540
                  +E    GL   S+  QD GN  AK + +  +++ +K+     Y ++    V
Sbjct: 577 AKE-----KGLTPPSTDHQD-SGNTEAKGAEAIYNRVKAAKKVPLDRMPYNLQ---YTVE 627

Query: 541 KEKNAIIYPHGDHHH 555
              +   ++I PH DH+H
Sbjct: 628 VKNGSLIIPHYDHYH 642


 Score = 66.6 bits (161), Expect = 2e-09
 Identities = 48/190 (25%), Positives = 80/190 (41%), Gaps = 30/190 (15%)

Query: 138 LKDAAHADNVRTKDEIN------RQKQEHVKDNEKV---NSNVAVARSQGRYTTNDGYVF 188
           LKD + +D V+ ++I      R  +   K N ++   + + VA+  G+YTT DGY+F
Sbjct: 478 LKDVS-SDKVKLVEDILAFLAPIRHPERLGKPNAQITYTDDEIQVAKLAGKYTTEDGYIF 536

Query: 189 NPADIIEDTGNAYIVPHGGHYHYIPXXXXXXXXXXXXXXXXXXXXXXXNMQPSQLSYSSTASDN 248
           +P DI  D G+AY+ PH  H H+I                        + P    ST   +
Sbjct: 537 DPRDITSDEGDAYVTPHMTHSHWIKKDSLSEAERAAAQAYAKEKGLTP-----PSTDHQD 591

Query: 249 NTQSVAKGSTSKPANKSENLQSLLKELYDSPSAQRYSESDGLVFDPAKIISRTPNGVAIP 308
            +  + AKG+                 + +Y+   A +    D + ++     +     + IP
Sbjct: 592 SGNTEAKGA--------------EAIYNRVKAAKKVPLDRMPYNLQYTVEVKNGSLIIP 636

Query: 309 HGDHYHFIPY 318
           H DHYH I +
Sbjct: 637 HYDHYHNIKF 646


 Score = 50.1 bits (118), Expect = 2e-04
 Identities = 55/302 (18%), Positives = 108/302 (35%), Gaps = 84/302 (27%)

Query: 297 IISRTPNGVAIPHGDHYHF----IPYSKLSALEEKIARRVPISGTGSTVSTNAKPNEVVX 352
           +I  T  G    HGDHYH+    +PY + +
Sbjct:  71 VIKITDQGYVTSHGDHYHYYNGKVPYDAIIS----------------------------- 101

Query: 353 XXXXXXXXXXXXXXXKELSSASDGYIFNPKDIVEETATAYIVR-HGDHFHYIPKS----- 406
                           +EL    Y     DIV E    Y+++ +G ++ Y+  +
Sbjct: 102 --------------EELLMKDPNYQLKDSDIVNEIKGGYVIKVNGKYYVYLKDAAHADN 146

Query: 407 --------NQIGQPTLPNNSLATPSPSLPINPGISHEKHEEDGYGFDANRIIAEDESGFI 458
                   Q  + +  +NS A + +      G      +DGY F+A+ II +    +I
Sbjct: 147 IRTKEEIKRQKQERSHNHNSRADNAVAAARAQG---RYTTDDGYIFNASDIIEDTGDAYI 203

Query: 459 MSHGNHNHYYFFKKDLTEEQIKAAQKHLEEVKTSHNGLDSLSSHEQDYPG---------- 507
            + HG+H HY  K +L+  ++ AA+ +     + S    S   +      P
Sbjct: 204 VPHGDHYHYIPKNELSASELAAAEAYWNGKQGSRPSSSSSYNANPAQPRLSENHNLTVTP 263

Query: 508 --NAKEMKDLDKKIEEKIAGIMKQYGVKRESI------VVNKEKNAIIYPHGDHHHADPI 559
             +  + +++   + E  A  + + V+ + +         + ++     + PHG+H+H  P
Sbjct: 264 TYHQNQGENISSLLRELYAKPLSERHVESDGLIFDPAQITSRTARGVAVPHGNHYHFIPY 323

Query: 560 DE 561
            ++
Sbjct: 324 EQ 325
```

Score = 44.7 bits (104), Expect = 0.009
Identities = 45/186 (24%), Positives = 63/186 (33%), Gaps = 36/186 (19%)

```
Query: 372  SASDGYIFNPKDIVEETATAYIVRHGDHFHYIPKSN------QIGQPTLPNNSLATPSPS 425
            +  DGYIF+P+DI  +   AY+ H  H H+I K +      Q      L  PS
Sbjct: 529  TTEDGYIFDPRDITSDEGDAYVTPHMTHSHWIKKDSLSEAERAAAQAYAKEKGLTPPSTD 588

Query: 426  LPINPGISHEKHEE--------------DGYGFDANRIIAEDESGFIMSHGNHNH----- 466
            + G + KE              D  ++  +    I+ H +H H
Sbjct: 589  HQ-DSGNTEAKGAEAIYNRVKAAKKVPLDRMPYNLQYTVEVKNGSLIIPHYDHYHNIKFE 647

Query: 467  ------YFFKKDLTEEQIKAAQK----HLEEVKTSHNGLDSLSSHEQDYPGNAKEMKDLD 516
                  Y  K TE+ A K     H E  SNG + SH Q      +    +
Sbjct: 648  WFDEGLYEAPKGYTLEDLLATVKYYVEHPNERPHSDNGFGNASDHVQRNKNGQADTNQTE 707

Query: 517  KKIEEK 522
            K  EEK
Sbjct: 708  KPSEEK 713
```

Score = 39.7 bits (91), Expect = 0.30
Identities = 34/138 (24%), Positives = 51/138 (36%), Gaps = 19/138 (13%)

```
Query: 274  ELYDSPSAQRYSESDGLVFDPAKIISRTPNGVAIPHGDHYHFIPYSKLSALEEKIARRV- 332
            E+ +  A +Y+  DG +FDP  I S   +   PH H H+I   LS  E   A+
Sbjct: 518  EIQVAKLAGKYTTEDGYIFDPRDITSDEGDAYVTPHMTHSHWIKKDSLSEAERAAAQAYA 577

Query: 333  -------PISGTGSTVSTNAKPNEVVXXXXXXXXXXXXXXXXXXXXKELSSASDGYIFNPKDIV 385
                   P +   + +T AK E +                   D +N + V
Sbjct: 578  KEKGLTPPSTDHQDSGNTEAKGAEAIYNRVKAAK----------KVPLDRMPYNLQYTV 626

Query: 386  EETATAYIVRHGDHFHYI 403
            E   + I+ H DH+H I
Sbjct: 627  EVKNGSLIIPHYDHYHNI 644
```

tr Q9ANY3 **Pneumococcal histidine triad protein B precursor (Fragment)**   819
          [phtB]   AA
          [Streptococcus pneumoniae]   align

Score = 423 bits (1087), Expect = e-116
Identities = 211/357 (59%), Positives = 271/357 (75%), Gaps = 11/357 (3%)

```
Query: 1    MKFSKKYIAAGSAVIVSLSLCAYALNQHRS-QENKDNNRVSYVDGSQSSQKSENLTPDQV 59
            MK +KKY+A GS  +++LS+C+Y L ++++  Q+ K++NRV+Y+DG Q+  QK+ENLTPD+V
Sbjct: 1    MKINKKYLA-GSVAVLALSVCSYELGRYQAGQDKKESNRVAYIDGDQAGQKAENLTPDEV 59

Query: 60   SQKEGIQAEQIVIKITDQGYVTSHGDHYHYYNGKVPYDALFSEELLMKDPNYQLKDADIV 119
            S++EGI AEQIVIKITDQGYVTSHGDHYHYYNGKVPYDA+ SEELLMKDPNYQLKD+DIV
Sbjct: 60   SKREGINAEQIVIKITDQGYVTSHGDHYHYYNGKVPYDAIISEELLMKDPNYQLKDSDIV 119

Query: 120  NEVKGGYIIKVDGKYYVYLKDAAHADNVRTKDEINRQKQEHVKD-NEKVNSNVAVARSQG 178
            NE+KGGY+IKV+GKYYVYLKDAAHADN+RTK+EI RQKQE  + N + ++ VA AR+QG
Sbjct: 120  NEIKGGYVIKVNGKYYVYLKDAAHADNIRTKEEIKRQKQERSHNHNSRADNAVAAARAQG 179
```

```
Query: 179  RYTTNDGYVFNPADIIEDTGNAYIVPHGGHYHYIPXXXXXXXXXXXXXXXXXXXXXXNMQPSQ 238
            RYTT+DGY+FN +DIIEDTG+AYIVPHG HYHYIP                        Q S+
Sbjct: 180  RYTTDDGYIFNASDIIEDTGDAYIVPHGDHYHYIP--KNELSASELAAAEAYWNGKQGSR 237

Query: 239  LSYSSTASDNNTQ---SVAKGSTSKPA---NKSENLQSLLKELYDSPSAQRYSESDGLVF 292
            S SS+ + N Q  S     T P   N+ EN+ SLL+ELY  P ++R+ ESDGL+F
Sbjct: 238  PSSSSSYNANPAQPRLSENHNLTVTPTYHQNQGENISSLLRELYAKPLSERHVESDGLIF 297

Query: 293  DPAKIISRTPNGVAIPHGDHYHFIPYSKLSALEEKIARRVPISGTGSTVSTNAKPNE 349
            DPA+I SRT  GVA+PHG+HYHFIPY ++S LE++IAR +P+    +   +++P E
Sbjct: 298  DPAQITSRTARGVAVPHGNHYHFIPYEQMSELEKRIARIIPLRYRSNHWVPDSRPEE 354


 Score = 75.9 bits (185), Expect = 4e-12
 Identities = 110/554 (19%), Positives = 208/554 (36%), Gaps = 93/554 (16%)

Query: 29   RSQENKDNNRVSYVDGSQSSQKSENLTPDQVSQKEG-IQAEQIVIKITDQGYVTSHGDHY 87
            R ++ + +N S D + ++ +++      + + +G I   +I+ T  Y+ HGDHY
Sbjct: 155  RQKQERSHNHNSRADNAVAAARAQG----RYTTDDGYIFNASDIIEDTGDAYIVPHGDHY 210

Query: 88   HYYNGKVPYDALFSEELLMKDPNYQLKDADIVNEVKGGYIIKVDGKYYVYLKDAAHADNV 147
            HY   +P + L + EL  +  + K  +            Y + A    +
Sbjct: 211  HY----IPKNELSASELAAAEAYWNGKQGSRPSSSSS-----------YNANPAQP-RL 253

Query: 148  RTKDEINRQKQEHVKDNEKVNSNVAVARSQG---RYTTNDGYVFNPADIIEDTGNAYIVP 204
                  +    H   E ++S +    ++    R+  +DG +F+PA I    T      VP
Sbjct: 254  SENHNLTVTPTYHQNQGENISSLLRELYAKPLSERHVESDGLIFDPAQITSRTARGVAVP 313

Query: 205  HGGHYHYIPXXXXXXXXXXXXXXXXXXXXXNMQPSQLSYSSTASDNNTQSVAKGSTSKPANK 264
            HG HYH+IP                     + P+  +   D+ +  +   T +P+
Sbjct: 314  HGNHYHFIPYEQMSELEKRIA-------RIIPLRYRSNHWVPDSRPEEPSPQPTPEPS-- 364

Query: 265  SENLQSLLKELYDSPSAQRYSESDGLVFDPAKIISRTPNGVAIPHGDHYHFIPYSKLSA- 323
                        SP    + DG +  A + + +G          +IP   LSA
Sbjct: 365  ------------PSPQPAPSNPIDGKLVKEA--VRKVGDGYVFEENGVSRYIPAKDLSAE 410

Query: 324  ----LEEKIARRVPISGTGSTVSTNAKP------NEVVXXXXXXXXXXXXXXXXXXKELSSA 373
                ++ K+A++  +S    T T+        N+                  +    A
Sbjct: 411  TAAGIDSKLAKQESLSHKLGTKKTDLPSSDREFYNKAYDLLARIHQDLLDNKGRQVDFEA 470

Query: 374  SDGYIFNPKDI------VEETATAYI--VRHGDHFHYIPKSNQIGQPTLPNNSLATPSPS 425
            D  +  KD+      + E  A++  +RH +       ++G+P   N  +
Sbjct: 471  LDNLLERLKDVSSDKVKLVEDILAFLAPIRHPE--------RLGKP---NAQITYTDDE 518

Query: 426  LPINPGISHEKHEEDGYGFDANRIIAEDESGFIMSHGNHNHYYFFKKDLTEEQIKAAQKHL 485
            + +   ++ +    EDGY FD  I +++   ++ H H+H+   K  L+E +  AAQ +
Sbjct: 519  IQVAK-LAGKYTAEDGYIFDPRDITSDEGDAYVTPHMTHSHWIKKDSLSEAERAAAQAYA 577

Query: 486  EEVKTSHNGLDSLSSHEQDYPGN--AKEMKDLDKKIE--EKIAGIMKQYGVKRESIVVNK 541
            E      GL   S+  QD GN  AK + +  +++  +K+      Y ++    V
Sbjct: 578  XE-----KGLTPPSTDHQD-SGNTEAKGAEAIYNRVKAAKKVPLDRMPYNLQ---YTVEV 628

Query: 542  EKNAIIYPHGDHHH 555
            +   ++I PH DH+H
Sbjct: 629  KNGSLIIPHYDHYH 642
```

```
Score = 73.9 bits (180), Expect = 1e-11
Identities = 66/268 (24%), Positives = 96/268 (35%), Gaps = 85/268 (31%)

Query:  297  IISRTPNGVAIPHGDHYHF----IPYSKLSALEEKIARRVP-----------ISGTGST 340
             +I  T  G   HGDHYH+     +PY  +  + E++  + P             I G G
Sbjct:   71  VIKITDQGYVTSHGDHYHYYNGKVPYDAI---ISEELLMKDPNYQLKDSDIVNEIKG-GYV 127

Query:  341  VSTNAK--------------PNEVVXXXXXXXXXXXXXXXXXXKELSSA-------SDGY 377
             +  N K                      E +                ++A        DGY
Sbjct:  128  IKVNGKYYVYLKDAAHADNIRTKEEIKRQKQERSHNHNSRADNAVAAARAQGRYTTDDGY 187

Query:  378  IFNPKDIVEETATAYIVRHGDHFHYIPKSN----------------------------Q 408
             IFN  DI+E+T  AYIV HGDH+HYIPK+
Sbjct:  188  IFNASDIIEDTGDAYIVPHGDHYHYIPKNELSASELAAAEAYWNGKQGSRPSSSSSYNAN 247

Query:  409  IGQPTLPNNSLATPSPSLPINPG--------------ISHEKHEEDGYGFDANRIIAEDE 454
              QP L N   T +P+  N G              +S    E DG  FD  +I +
Sbjct:  248  PAQPRLSENHNLTVTPTYHQNQGENISSLLRELYAKPLSERHVESDGLIFDPAQITSRTA 307

Query:  455  SGFIMSHGNHNHYFFKKDLTEEQIKAAQ 482
              G  + HGNH H+   + ++E + + A+
Sbjct:  308  RGVAVPHGNHYHFIPYEQMSELEKRIAR 335


Score = 64.7 bits (156), Expect = 9e-09
Identities = 48/190 (25%), Positives = 80/190 (41%), Gaps = 30/190 (15%)

Query:  138  LKDAAHADNVRTKDEIN------RQKQEHVKDNEKV---NSNVAVARSQGRYTTNDGYVF 188
             LKD + +D V+  ++I      R  +   K N ++   +  + VA+  G+YT  DGY+F
Sbjct:  478  LKDVS-SDKVKLVEDILAFLAPIRHPERLGKPNAQITYTDDEIQVAKLAGKYTAEDGYIF 536

Query:  189  NPADIIEDTGNAYIVPHGGHYHYIPXXXXXXXXXXXXXXXXXXXXXNMQPSQLSYSSTASDN 248
             +P DI  D G+AY+  PH  H H+I                      + P   ST   +
Sbjct:  537  DPRDITSDEGDAYVTPHMTHSHWIKKDSLSEAERAAAQAYAXEKGLTP-----PSTDHQD 591

Query:  249  NTQSVAKGSTSKPANKSENLQSLLKELYDSPSAQRYSESDGLVFDPAKIISRTPNGVAIP 308
             +  + AKG+                 + +Y+  A +   D + ++    +      + IP
Sbjct:  592  SGNTEAKGA---------------EAIYNRVKAAKKVPLDRMPYNLQYTVEVKNGSLIIP 636

Query:  309  HGDHYHFIPY 318
             H DHYH I +
Sbjct:  637  HYDHYHNIKF 646


Score = 47.4 bits (111), Expect = 0.001
Identities = 47/255 (18%), Positives = 92/255 (35%), Gaps = 77/255 (30%)

Query:  384  IVEETATAYIVRHGDHFHY----IPKSNQIGQPTL---PNNSL---------------- 419
             +++ T  Y+  HGDH+HY     +P    I +  L   PN  L
Sbjct:   71  VIKITDQGYVTSHGDHYHYYNGKVPYDAIISEELLMKDPNYQLKDSDIVNEIKGGYVIKV 130

Query:  420  ----------ATPSPSLPINPGISHEKHE----------------------EDGYGFD 445
                       A + ++     I +K E                        +DGY F+
Sbjct:  131  NGKYYVYLKDAAHADNIRTKEEIKRQKQERSHNHNSRADNAVAAARAQGRYTTDDGYIFN 190

Query:  446  ANRIIAEDESGFIMSHGNHNHYFFKKDLTEEQIKAAQKHLEEVKTSHNGLDSLSSHEQDY 505
             A+ II +    +I+ HG+H HY  K +L+  ++ AA+ +    + S    S   S +
Sbjct:  191  ASDIIEDTGDAYIVPHGDHYHYIPKNELSASELAAAEAYWNGKQGSRPSSSSSYNANPAQ 250
```

```
Query: 506  PG-------------NAKEMKDLDKKIEEKIAGIMKQYGVKRESI------VVNKEKNAI 546
            P              +  +  +++   +  E  A  ++  V+ ++      +  ++     +
Sbjct: 251  PRLSENHNLTVTPTYHQNQGENISSLLRELYAKPLSERHVESDGLIFDPAQITSRTARGV 310

Query: 547  IYPHGDHHHADPIDE 561
                PHG+H+H   P  ++
Sbjct: 311  AVPHGNHYHFIPYEQ 325
```

```
 Score = 45.8 bits (107), Expect = 0.004
 Identities = 45/185 (24%), Positives = 65/185 (34%), Gaps = 34/185 (18%)
```

```
Query: 372  SASDGYIFNPKDIVEETATAYIVRHGDHFHYIPKSNQIGQPTLPNNSLA-----TPSPSL 426
            +A DGYIF+P+DI  +   AY+  H  H  H+I K  +           +  A    TP  +
Sbjct: 529  TAEDGYIFDPRDITSDEGDAYVTPHMTHSHWIKKDSLSEAERAAAQAYAXEKGLTPPSTD 588

Query: 427  PINPGISHEKHEE--------------DGYGFDANRIIAEDESGFIMSHGNHNH------ 466
            +  G +  K  E              D   ++     +        I+ H +H H
Sbjct: 589  HQDSGNTEAKGAEAIYNRVKAAKKVPLDRMPYNLQYTVEVKNGSLIIPHYDHYHNIKFEW 648

Query: 467  -----YFFKKDLTEEQIKAAQK----HLEEVKTSHNGLDSLSSHEQDYPGNAKEMKDLDK 517
                 Y   K  TE +A K       H E    SNG  + S HQ        +     +K
Sbjct: 649  FDEGLYEAPKGYTLEDLLATVKYYVEHPNERPHSDNGFGNASDHVQRNKNGQADTNQTEK 708

Query: 518  KIEEK 522
            EEK
Sbjct: 709  PSEEK 713
```

```
 Score = 40.0 bits (92), Expect = 0.23
 Identities = 32/130 (24%), Positives = 49/130 (37%), Gaps = 3/130 (2%)
```

```
Query: 274  ELYDSPSAQRYSESDGLVFDPAKIISRTPNGVAIPHGDHYHFIPYSKLSALEEKIARRVP 333
            E+  +  A +Y+  DG +FDP  I S    +    PH  H H+I    LS  E    A+
Sbjct: 518  EIQVAKLAGKYTAEDGYIFDPRDITSDEGDAYVTPHMTHSHWIKKDSLSEAERAAAQAYA 577

Query: 334  ISGTGSTVSTNAKPNEVVXXXXXXXXXXXXXXXXXXXKELSSASDGYIFNPKDIVEETATAYI 393
            +   ST+ + +                         K      D +N +  VE   + I
Sbjct: 578  XEKGLTPPSTDHQDSGNTEAKGAEAIYNRVKAAKK---VPLDRMPYNLQYTVEVKNGSLI 634

Query: 394  VRHGDHFHYI 403
            +  H DH+H I
Sbjct: 635  IPHYDHYHNI 644
```

```
tr Q6WNQ3  Surface protein BVH-11 (Fragment) [bvh-11] [Streptococcus  811 AA
           pneumoniae]
                                                                     align
```

```
 Score =  417 bits (1071), Expect = e-115
 Identities = 208/337 (61%), Positives = 256/337 (75%), Gaps = 10/337 (2%)
```

```
Query: 21   CAYALNQHRS-QENKDNNRVSYVDGSQSSQKSENLTPDQVSQKEGIQAEQIVIKITDQGY 79
            C+Y L +H++ Q  K++NRVSY+DG Q+ QK+ENLTPD+VS++EGI AEQIVIKITDQGY
Sbjct: 1    CSYELGRHQAGQVKKESNRVSYIDGDQAGQKAENLTPDEVSKREGINAEQIVIKITDQGY 60
```

```
Query:  80  VTSHGDHYHYYNGKVPYDALFSEELLMKDPNYQLKDADIVNEVKGGYIIKVDGKYYVYLK 139
            VTSHGDHYHYYNGKVPYDA+ SEELLMKDPNYQLKD+DIVNE+KGGY+IKVDGKYYVYLK
Sbjct:  61  VTSHGDHYHYYNGKVPYDAIISEELLMKDPNYQLKDSDIVNEIKGGYVIKVDGKYYVYLK 120

Query: 140  DAAHADNVRTKDEINRQKQEHVKD-NEKVNSNVAVARSQGRYTTNDGYVFNPADIIEDTG 198
            DAAHADN+RTK+EI RQKQE    + N + ++ VA AR+QGRYTT+DGY+FN +DIIEDTG
Sbjct: 121  DAAHADNIRTKEEIKRQKQERSHNHNSRADNAVAAARAQGRYTTDDGYIFNASDIIEDTG 180

Query: 199  NAYIVPHGGHYHYIPXXXXXXXXXXXXXXXXXXXXXXXNMQPSQLSYSSTASDNNTQ---SVAK 255
            +AYIVPHG HYHYIP                       Q S+ S SS+ + N  Q     S
Sbjct: 181  DAYIVPHGDHYHYIP--KSDLSASELAAAQAYWNGKQGSRPSSSSSHNANPAQPRLSENH 238

Query: 256  GSTSKPA---NKSENLQSLLKELYDSPSAQRYSESDGLVFDPAKIISRTPNGVAIPHGDH 312
                 T  P    N+ EN+ SLL+ELY  P ++R+ ESDGL+FDPA+I SRT NGVA+PHGDH
Sbjct: 239  NLTVTPTYHQNQGENISSLLRELYAKPLSERHVESDGLIFDPAQITSRTANGVAVPHGDH 298

Query: 313  YHFIPYSKLSALEEKIARRVPISGTGSTVSTNAKPNE 349
            YHFIPYS+LS LEEK+AR +P+    +    +++P +
Sbjct: 299  YHFIPYSQLSPLEEKLARIIPLRYRSNHWVPDSRPEQ 335


  Score = 84.3 bits (207), Expect = 1e-14
  Identities = 90/385 (23%), Positives = 134/385 (34%), Gaps = 73/385 (18%)

Query: 193  IIEDTGNAYIVPHGGHYHYI-----------------PXXXXXXXXXXXXXXXXXXXXNM 234
            +I+ T   Y+  HG HYHY                  P                     +
Sbjct:  52  VIKITDQGYVTSHGDHYHYYNGKVPYDAIISEELLMKDPNYQLKDSDIVNEIKGGYVIKV 111

Query: 235  QPSQLSYSSTASDNNTQSVAKGSTSKPANKSENLQSLLKELYDSPSAQ-RYSESDGLVFD 293
                 Y  A+ +   +   +   +S N  S        + AQ RY+  DG +F+
Sbjct: 112  DGKYYVYLKDAAHADNIRTKEEIKRQKQERSHNHNSRADNAVAAARAQGRYTTDDGYIFN 171

Query: 294  PAKIISRTPNGVAIPHGDHYHFIPYSKLSALEEKIARRVPISGTGSTVST----NAKP-- 347
            + II  T +    +PHGDHYH+IP S LSA E   A+        GS  S+    NA P
Sbjct: 172  ASDIIEDTGDAYIVPHGDHYHYIPKSDLSASELAAAQAYWNGKQGSRPSSSSSHNANPAQ 231

Query: 348  -------NEVVXXXXXXXXXXXXXXXXXXKELSSA--------SDGYIFNPKDIVEETATAY 392
                   N V                   +EL +         SDG IF+P I  TA
Sbjct: 232  PRLSENHNLTVTPTYHQNQGENISSLLRELYAKPLSERHVESDGLIFDPAQITSRTANGV 291

Query: 393  IVRHGDHFHYIPKS----------------------NQIGQPTLPNNSLATPSPSLP 427
            V HGDH+H+IP S                       ++  QP+ +     +PSP
Sbjct: 292  AVPHGDHYHFIPYSQLSPLEEKLARIIPLRYRSNHWVPDSRPEQPSPQSTPEPSPSPQPA 351

Query: 428  INPGISHEKHEEDGYGFDANRIIAEDESGFIMSHGNHNHYYFFKKDLTEEQIKAAQKHL-E 486
            +NP +     ++    +A R + +   G++        Y  KDL+ E          L +
Sbjct: 352  PNPQPAPSNPIDEKLVKEAVRKVGD---GYVFEENGVPRYIPAKDLSAETAAGIDSKLAK 408

Query: 487  EVKTSHNGLDSLSSHEQDYPGNAKE 511
            +   SH    L + + D P + +E
Sbjct: 409  QESLSH----KLGAKKTDLPSSDRE 429


  Score = 78.2 bits (191), Expect = 8e-13
  Identities = 112/541 (20%), Positives = 205/541 (37%), Gaps = 61/541 (11%)

Query:  29  RSQENKDNNRVSYVDGSQSSQKSENLTPDQVSQKEG-IQAEQIVIKITDQGYVTSHGDHY 87
```

```
                 R ++ + +N  S  D + ++ +++     + +  +G I     +I+ T   Y+   HGDHY
Sbjct:  136  RQKQERSHNHNSRADNAVAAARAQG----RYTTDDGYIFNASDIIEDTGDAYIVPHGDHY  191

Query:   88  HYYNGKVPYDALFSEELLMKDPNYQLKDADIVNEVKGGYIIKVDGKYYVYLKDAAHADNV  147
             HY   +P  L + EL     + K   +           +         +     + N+
Sbjct:  192  HY----IPKSDLSASELAAAQAYWNGKQGSRPSSSSSHNANPAQPRL-------SENHNL  240

Query:  148  RTKDEINRQKQEHVKDNEKVNSNVAVARSQGRYTTNDGYVFNPADIIEDTGNAYIVPHGG  207
                 ++ + E++ +  +    A   S+ R+ +DG +F+PA I   T N   VPHG
Sbjct:  241  TVTPTYHQNQGENI--SSLLRELYAKPLSE-RHVESDGLIFDPAQITSRTANGVAVPHGD  297

Query:  208  HYHYIPXXXXXXXXXXXXXXXXXXXXX-------NMQPSQLSYSSTASDNNTQSVAKGSTSK  260
             HYH+IP                          + +P Q S  ST   + +     A
Sbjct:  298  HYHFIPYSQLSPLEEKLARIIPLRYRSNHWVPDSRPEQPSPQSTPEPSPSPQPAPNPQPA  357

Query:  261  PANKSENLQSLLKELYDSPSAQRYSESDGLV-FDPAKIIS-RTPNGVAIPHGDHYHFIPY  318
             P+N +  + L+KE        E +G+  + PAK +S  T G+
Sbjct:  358  PSNPID--EKLVKEAVRKVGDGYVFEENGVPRYIPAKDLSAETAAGID-----------  403

Query:  319  SKLSALEEKIARRVPISGTGSTVSTNAKPNEVVXXXXXXXXXXXXXXXXXXKELSSASDGYI  378
             SKL A +E ++ ++    T   S   N+                 +    A D +
Sbjct:  404  SKL-AKQESLSHKLGAKKTDLPSSDREFYNKAYDLLARIHQDLLDNKGRQVDFEALDNLL  462

Query:  379  FNPKDIVEETATAYIVRHGDHFHYIPKSNQIGQPTLPNNSLATPSPSLPINPGISHEKHE  438
               KD+ +        D   ++          +    PN +       + +     ++ +
Sbjct:  463  ERLKDVSSDKVKLV----DDILAFLAPIRHPERLGKPNAQITYTDDEIQVAK-LAGKYTT  517

Query:  439  EDGYGFDANRIIAEDESGFIMSHGNHNHYFFKKDLTEEQIKAAQKHLEEVKTSHNGLDSL  498
             EDGY FD   I +++    ++  H  H+H+  K  L+E +  AAQ + +E      GL
Sbjct:  518  EDGYIFDPRDITSDEGDAYVTPHMTHSHWIKKDSLSEAERAAAQAYAKE-----KGLTPP  572

Query:  499  SSHEQDYPGN--AKEMKDLDKKIE--EKIAGIMKQYGVKRESIVVNKEKNAIIYPHGDHH  554
             S+   QD  GN   AK  + +  +++     +K+      Y ++    V +  ++I PH DH+
Sbjct:  573  STDHQD-SGNTEAKGAEAIYNRVKAAKKVPLDRMPYNLQ---YTVEVKNGSLIIPHYDHY  628

Query:  555  H  555
             H
Sbjct:  629  H  629
```

```
 Score = 68.2 bits (165), Expect = 8e-10
 Identities = 49/190 (25%), Positives = 80/190 (41%), Gaps = 30/190 (15%)
```

```
Query:  138  LKDAAHADNVRTKDEIN------RQKQEHVKDNEKV---NSNVAVARSQGRYTTNDGYVF  188
             LKD + +D V+ D+I       R  +  K N ++   +  + VA+  G+YTT DGY+F
Sbjct:  465  LKDVS-SDKVKLVDDILAFLAPIRHPERLGKPNAQITYTDDEIQVAKLAGKYTTEDGYIF  523

Query:  189  NPADIIEDTGNAYIVPHGGHYHYIPXXXXXXXXXXXXXXXXXXXXXXNMQPSQLSYSSTASDN  248
             +P DI  D G+AY+ PH  H H H+I                    + P    ST   +
Sbjct:  524  DPRDITSDEGDAYVTPHMTHSHWIKKDSLSEAERAAAQAYAKEKGLTP-----PSTDHQD  578

Query:  249  NTQSVAKGSTSKPANKSENLQSLLKELYDSPSAQRYSESDGLVFDPAKIISRTPNGVAIP  308
             +   + AKG+           + +Y+  A +   D + ++      +      + IP
Sbjct:  579  SGNTEAKGA--------------EAIYNRVKAAKKVPLDRMPYNLQYTVEVKNGSLIIP  623

Query:  309  HGDHYHFIPY  318
             H DHYH I +
Sbjct:  624  HYDHYHNIKF  633
```

```
Score = 58.2 bits (139), Expect = 8e-07
Identities = 60/306 (19%), Positives = 112/306 (35%), Gaps = 84/306 (27%)

Query: 297  IISRTPNGVAIPHGDHYHF----IPYSKLSALEEKIARRVPISGTGSTVSTNAKPNEVVX 352
            +I   T  G    HGDHYH+     +PY  + +
Sbjct: 52   VIKITDQGYVTSHGDHYHYYNGKVPYDAIIS-------------------------- 82

Query: 353  XXXXXXXXXXXXXXXXXKELSSASDGYIFNPKDIVEETATAYIVR-HGDHFHYIPKS----- 406
                             +EL    Y     DIV E     Y+++  G ++ Y+  +
Sbjct: 83   ---------------EELLMKDPNYQLKDSDIVNEIKGGYVIKVDGKYVVYLKDAAHADN 127

Query: 407  --------NQIGQPTLPNNSLATPSPSLPINPGISHEKHEEDGYGFDANRIIAEDESGFI 458
                    Q  + +  +NS A + +     G       +DGY F+A+ II +    +I
Sbjct: 128  IRTKEEIKRQKQERSHNHNSRADNAVAAARAQG---RYTTDDGYIFNASDIIEDTGDAYI 184

Query: 459  MSHGNHNHYFFKKDLTEEQIKAAQKHLE-------EVKTSHNGLDSLSSHEQDY------ 505
            + HG+H HY  K DL+  ++ AAQ +        +SHN   +      +++
Sbjct: 185  VPHGDHYHYIPKSDLSASELAAAQAYWNGKQGSRPSSSSSHNANPAQPRLSENHNLTVTP 244

Query: 506  PGNAKEMKDLDKKIEEKIAGIMKQYGVKRESI------VVNKEKNAIIYPHGDHHHADPI 559
             +  + +++   + E  A + +  V+ + +       + ++  N +   PHGDH+H  P
Sbjct: 245  TYHQNQGENISSLLRELYAKPLSERHVESDGLIFDPAQITSRTANGVAVPHGDHYHFIPY 304

Query: 560  DEHKPV 565
             +  P+
Sbjct: 305  SQLSPL 310


Score = 42.7 bits (99), Expect = 0.036
Identities = 44/186 (23%), Positives = 63/186 (33%), Gaps = 36/186 (19%)

Query: 372  SASDGYIFNPKDIVEETATAYIVRHGDHFHYIPKSN------QIGQPTLPNNSLATPSPS 425
            +  DGYIF+P+DI  +   AY+  H  H H+I K +       Q       L  PS
Sbjct: 516  TTEDGYIFDPRDITSDEGDAYVTPHMTHSHWIKKDSLSEAERAAAQAYAKEKGLTPPSTD 575

Query: 426  LPINPGISHEKHEE--------------DGYGFDANRIIAEDESGFIMSHGNHNH----- 466
            + G + K E              D  ++     +       I+ H +H H
Sbjct: 576  HQ-DSGNTEAKGAEAIYNRVKAAKKVPLDRMPYNLQYTVEVKNGSLIIPHYDHYHNIKFE 634

Query: 467  ------YFFKKDLTEEQIKAAQK----HLEEVKTSHNGLDSLSSHEQDYPGNAKEMKDLD 516
                  Y  K  +E+ A K     H  E   S NG  + S H Q       +     +
Sbjct: 635  WFDEGLYEAPKGYSLEDLLATVKYYVEHPNERPHSDNGFGNASDHVQRNKNGQADTNQTE 694

Query: 517  KKIEEK 522
            K  EEK
Sbjct: 695  KPNEEK 700


Score = 39.7 bits (91), Expect = 0.30
Identities = 34/138 (24%), Positives = 51/138 (36%), Gaps = 19/138 (13%)

Query: 274  ELYDSPSAQRYSESDGLVFDPAKIISRTPNGVAIPHGDHYHFIPYSKLSALEEKIARRV- 332
            E+  + A +Y+  DG +FDP  I S   +   PH   H H+I   LS E   A+
Sbjct: 505  EIQVAKLAGKYTTEDGYIFDPRDITSDEGDAYVTPHMTHSHWIKKDSLSEAERAAAQAYA 564

Query: 333  -------PISGTGSTVSTNAKPNEVVXXXXXXXXXXXXXXXXXXXXXKELSSASDGYIFNPKDIV 385
```

```
                       P +      + +T AK  E +                          D  +N +  V
Sbjct: 565 KEKGLTPPSTDHQDSGNTEAKGAEAIYNRVKAAK-----------KVPLDRMPYNLQYTV 613


Query: 386 EETATAYIVRHGDHFHYI 403
               E    + I+ H DH+H I
Sbjct: 614 EVKNGSLIIPHYDHYHNI 631
```

tr Q6WNP8 **Surface protein BVH-11-2 (Fragment) [bvh-11-2] [Streptococcus 834 AA
             pneumoniae]**

align


```
 Score =  417 bits (1071), Expect = e-115
 Identities = 208/337 (61%), Positives = 256/337 (75%), Gaps = 10/337 (2%)

Query: 21  CAYALNQHRS-QENKDNNRVSYVDGSQSSQKSENLTPDQVSQKEGIQAEQIVIKITDQGY 79
           C+Y L +H++ Q  K++NRVSY+DG Q+ QK+ENLTPD+VS++EGI AEQIVIKITDQGY
Sbjct: 1   CSYELGRHQAGQVKKESNRVSYIDGDQAGQKAENLTPDEVSKREGINAEQIVIKITDQGY 60

Query: 80  VTSHGDHYHYYNGKVPYDALFSEELLMKDPNYQLKDADIVNEVKGGYIIKVDGKYYVYLK 139
           VTSHGDHYHYYNGKVPYDA+ SEELLMKDPNYQLKD+DIVNE+KGGY+IKVDGKYYVYLK
Sbjct: 61  VTSHGDHYHYYNGKVPYDAIISEELLMKDPNYQLKDSDIVNEIKGGYVIKVDGKYYVYLK 120

Query: 140 DAAHADNVRTKDEINRQKQEHVKD-NEKVNSNVAVARSQGRYTTNDGYVFNPADIIEDTG 198
           DAAHADN+RTK+EI RQKQE   + N + ++ VA AR+QGRYTT+DGY+FN +DIIEDTG
Sbjct: 121 DAAHADNIRTKEEIKRQKQERSHNHNSRADNAVAAARAQGRYTTDDGYIFNASDIIEDTG 180

Query: 199 NAYIVPHGGHYHYIPXXXXXXXXXXXXXXXXXXXXNMQPSQLSYSSTASDNNTQ---SVAK 255
           +AYIVPHG HYHYIP                      Q S+ S SS+ + N  Q     S
Sbjct: 181 DAYIVPHGDHYHYIP--KSDLSASELAAAQAYWNGKQGSRPSSSSSHNANPAQPRLSENH 238

Query: 256 GSTSKPA---NKSENLQSLLKELYDSPSAQRYSESDGLVFDPAKIISRTPNGVAIPHGDH 312
              T P    N+ EN+ SLL+ELY  P ++R+ ESDGL+FDPA+I SRT NGVA+PHGDH
Sbjct: 239 NLTVTPTYHQNQGENISSLLRELYAKPLSERHVESDGLIFDPAQITSRTANGVAVPHGDH 298

Query: 313 YHFIPYSKLSALEEKIARRVPISGTGSTVSTNAKPNE 349
           YHFIPYS+LS LEEK+AR +P+      +   +++P +
Sbjct: 299 YHFIPYSQLSPLEEKLARIIPLRYRSNHWVPDSRPEQ 335


 Score = 84.3 bits (207), Expect = 1e-14
 Identities = 90/385 (23%), Positives = 134/385 (34%), Gaps = 73/385 (18%)

Query: 193 IIEDTGNAYIVPHGGHYHYI------------------PXXXXXXXXXXXXXXXXXXXXNM 234
           +I+ T  Y+  HG HYHY                    P                     +
Sbjct: 52  VIKITDQGYVTSHGDHYHYYNGKVPYDAIISEELLMKDPNYQLKDSDIVNEIKGGYVIKV 111

Query: 235 QPSQLSYSSTASDNNTQSVAKGSTSKPANKSENLQSLLKELYDSPSAQ-RYSESDGLVFD 293
                  Y   A+  +     +     + +S N S    + AQ RY+  DG +F+
Sbjct: 112 DGKYYVYLKDAAHADNIRTKEEIKRQKQERSHNHNSRADNAVAAARAQGRYTTDDGYIFN 171

Query: 294 PAKIISRTPNGVAIPHGDHYHFIPYSKLSALEEKIARRVPISGTGSTVST----NAKP-- 347
           + II  T +  +PHGDHYH+IP S LSA E  A+    GS  S+    NA P
Sbjct: 172 ASDIIEDTGDAYIVPHGDHYHYIPKSDLSASELAAAQAYWNGKQGSRPSSSSSHNANPAQ 231

Query: 348 -------NEVVXXXXXXXXXXXXXXXXXXXKELSSA--------SDGYIFNPKDIVEETATAY 392
```

```
                       N  V                    +EL +           SDG IF+P  I    TA
Sbjct: 232 PRLSENHNLTVTPTYHQNQGENISSLLRELYAKPLSERHVESDGLIFDPAQITSRTANGV 291

Query: 393 IVRHGDHFHYIPKS--------------------------NQIGQPTLPNNSLATPSPSLP 427
           V HGDH+H+IP S                          ++ QP+ +     +PSP
Sbjct: 292 AVPHGDHYHFIPYSQLSPLEEKLARIIPLRYRSNHWVPDSRPEQPSPQSTPEPSPSPQPA 351

Query: 428 INPGISHEKHEEDGYGFDANRIIAEDESGFIMSHGNHNHYFFKKDLTEEQIKAAQKHL-E 486
           NP +         ++    +A R + +  G++       Y  KDL+ E        L +
Sbjct: 352 PNPQPAPSNPIDEKLVKEAVRKVGD---GYVFEENGVPRYIPAKDLSAETAAGIDSKLAK 408

Query: 487 EVKTSHNGLDSLSSHEQDYPGNAKE 511
           +   SH     L + + D P + +E
Sbjct: 409 QESLSH----KLGAKKTDLPSSDRE 429


 Score = 68.2 bits (165), Expect = 8e-10
 Identities = 49/190 (25%), Positives = 80/190 (41%), Gaps = 30/190 (15%)

Query: 138 LKDAAHADNVRTKDEIN------RQKQEHVKDNEKV---NSNVAVARSQGRYTTNDGYVF 188
           LKD + +D V+ D+I       R  +  K N ++   +  + VA+ G+YTT DGY+F
Sbjct: 465 LKDVS-SDKVKLVDDILAFLAPIRHPERLGKPNAQITYTDDEIQVAKLAGKYTTEDGYIF 523

Query: 189 NPADIIEDTGNAYIVPHGGHYHYIPXXXXXXXXXXXXXXXXXXXXXXXNMQPSQLSYSSTASDN 248
           +P DI  D G+AY+ PH  H H+I                        + P    ST   +
Sbjct: 524 DPRDITSDEGDAYVTPHMTHSHWIKKDSLSEAERAAAQAYAKEKGLTP-----PSTDHQD 578

Query: 249 NTQSVAKGSTSKPANKSENLQSLLKELYDSPSAQRYSESDGLVFDPAKIISRTPNGVAIP 308
           + + +AKG+               + +Y+   A +    D + ++      +       + IP
Sbjct: 579 SGNTEAKGA---------------EAIYNRVKAAKKVPLDRMPYNLQYTVEVKNGSLIIP 623

Query: 309 HGDHYHFIPY 318
           H DHYH I +
Sbjct: 624 HYDHYHNIKF 633


 Score = 58.2 bits (139), Expect = 8e-07
 Identities = 60/306 (19%), Positives = 112/306 (35%), Gaps = 84/306 (27%)

Query: 297 IISRTPNGVAIPHGDHYHF----IPYSKLSALEEKIARRVPISGTGSTVSTNAKPNEVVX 352
           +I  T  G     HGDHYH+      +PY + +
Sbjct: 52  VIKITDQGYVTSHGDHYHYYNGKVPYDAIIS----------------------------- 82

Query: 353 XXXXXXXXXXXXXXXXKELSSASDGYIFNPKDIVEETATAYIVR-HGDHFHYIPKS----- 406
                           +EL    Y    DIV E    Y+++  G ++ Y+   +
Sbjct: 83  --------------EELLMKDPNYQLKDSDIVNEIKGGYVIKVDGKYYVYLKDAAHADN 127

Query: 407 --------NQIGQPTLPNNSLATPSPSLPINPGISHEKHEEDGYGFDANRIIAEDESGFI 458
                   Q + +  +NS A + +     G        +DGY F+A+ II +    +I
Sbjct: 128 IRTKEEIKRQKQERSHNHNSRADNAVAAARAQG---RYTTDDGYIFNASDIIEDTGDAYI 184

Query: 459 MSHGNHNHYFFKKDLTEEQIKAAQKHLE-------EVKTSHNGLDSLSSHEQDY------ 505
           + HG+H HY K DL+ ++ AAQ +             +SHN +        +++
Sbjct: 185 VPHGDHYHYIPKSDLSASELAAAQAYWNGKQGSRPSSSSSHNANPAQPRLSENHNLTVTP 244

Query: 506 PGNAKEMKDLDKKIEEKIAGIMKQYGVKRESI------VVNKEKNAIIYPHGDHHHADPI 559
           + +  +++   + + E  A  + + +         + ++ N  + PHGDH+H  P
Sbjct: 245 TYHQNQGENISSLLRELYAKPLSERHVESDGLIFDPAQITSRTANGVAVPHGDHYHFIPY 304
```

```
Query: 560 DEHKPV 565
                +  P+
Sbjct: 305 SQLSPL 310
```

```
 Score = 47.4 bits (111), Expect = 0.001
 Identities = 61/260 (23%), Positives = 90/260 (34%), Gaps = 51/260 (19%)

Query: 372 SASDGYIFNPKDIVEETATAYIVRHGDHFHYIPKSN------QIGQPTLPNNSLATPSPS 425
            +  DGYIF+P+DI  +    AY+  H  H H+I K +       Q      L  PS
Sbjct: 516 TTEDGYIFDPRDITSDEGDAYVTPHMTHSHWIKKDSLSEAERAAAQAYAKEKGLTPPSTD 575

Query: 426 LPINPGISHEKHEE--------------DGYGFDANRIIAEDESGFIMSHGNHNH----- 466
            + G + K E                 D  ++   +    I+ H +H H
Sbjct: 576 HQ-DSGNTEAKGAEAIYNRVKAAKKVPLDRMPYNLQYTVEVKNGSLIIPHYDHYHNIKFE 634

Query: 467 ------YFFKKDLTEEQIKAAQK----HLEEVKTSHNGLDSLSSHEQDYPGNAKEMKDLD 516
                  Y   K  +E+ A  K     H  E  S NG  + S H Q       +     +
Sbjct: 635 WFDEGLYEAPKGYSLEDLLATVKYYVEHPNERPHSDNGFGNASDHVQRNKNGQADTNQTE 694

Query: 517 KKIEEKIAGIMKQYGVKRESIVVNKEKNAIIYPHGDHHHADPIDEHKPVGIGHSHSNYEL 576
            K  EEK        + E   +KE + +  P  H +D + H  VG+  S  N   L
Sbjct: 695 KPNEEK--------PQTEKPEEDKEHDEVSEP--THPESDEKENH--VGLNPSADN--L 739

Query: 577 FKPEEGVAKKEGNKVYTGEE 596
            +KP      + E     T +E
Sbjct: 740 YKPSTDTEETEEEAEDTTDE 759
```

```
 Score = 39.7 bits (91), Expect = 0.30
 Identities = 34/138 (24%), Positives = 51/138 (36%), Gaps = 19/138 (13%)

Query: 274 ELYDSPSAQRYSESDGLVFDPAKIISRTPNGVAIPHGDHYHFIPYSKLSALEEKIARRV- 332
            E+   + A +Y+  DG +FDP  I S   +   PH  H H+I    LS  E   A+
Sbjct: 505 EIQVAKLAGKYTTEDGYIFDPRDITSDEGDAYVTPHMTHSHWIKKDSLSEAERAAAQAYA 564

Query: 333 -------PISGTGSTVSTNAKPNEVVXXXXXXXXXXXXXXXXXXKELSSASDGYIFNPKDIV 385
                   P +   + +T AK  E +                    D  +N + V
Sbjct: 565 KEKGLTPPSTDHQDSGNTEAKGAEAIYNRVKAAK----------KVPLDRMPYNLQYTV 613

Query: 386 EETATAYIVRHGDHFHYI 403
            E    + I+ H DH+H I
Sbjct: 614 EVKNGSLIIPHYDHYHNI 631
```

tr <u>Q6WNQ1</u> **Surface protein BVH-11 (Fragment) [bvh-11] [Streptococcus** 825 AA
                pneumoniae]

<u>align</u>

```
 Score =  415 bits (1066), Expect = e-114
 Identities = 206/350 (58%), Positives = 257/350 (72%), Gaps = 21/350 (6%)

Query: 21 CAYALNQHRSQENKDNNRVSYVDGSQSSQKSENLTPDQVSQKEGIQAEQIVIKITDQGYV 80
           C+Y L ++++  K+NNRVSY+DG Q++QK+ENLTPD+VS++EGI AEQIVIKITDQGYV
Sbjct: 1  CSYELGLYQARPVKENNRVSYIDGKQATQKTENLTPDEVSKREGINAEQIVIKITDQGYV 60
```

```
Query:  81  TSHGDHYHYYNGKVPYDALFSEELLMKDPNYQLKDADIVNEVKGGYIIKVDGKYYVYLKD  140
            TSHGDHYHYYNGKVPYDA+FSEELLMKDPNY+LKD DIVNEVKGGY+IKVDGKYYVYLKD
Sbjct:  61  TSHGDHYHYYNGKVPYDAIFSEELLMKDPNYKLKDEDIVNEVKGGYVIKVDGKYYVYLKD  120

Query: 141  AAHADNVRTKDEINRQKQEHVKDNE----KVNSNVAVARSQGRYTTNDGYVFNPADIIED  196
            AAHADNVRTK+EINRQKQEH + E   + + VA+ARSQGRYTT+DGY+FN +DIIED
Sbjct: 121  AAHADNVRTKEEINRQKQEHSQHREGGTPRNDGAVALARSQGRYTTDDGYIFNASDIIED  180

Query: 197  TGNAYIVPHGGHYHYIPXXXXXXXXXXXXXXXXXXXXXNMQPSQLSYSSTASDN--------  248
            TG+AYIVPHG HYHYIP                  +  +Y    SDN
Sbjct: 181  TGDAYIVPHGDHYHYIPKNELSASELAAAEAFLSGRGNLSNSRTYRLQNSDNTPRTNWVP  240

Query: 249  --------NTQSVAKGSTSKPANKS-ENLQSLLKELYDSPSAQRYSESDGLVFDPAKIIS  299
                    NT +    +T+  A++S E++ SLLK+LY  P +QR+ ESDGL+FDPA+I S
Sbjct: 241  SVSNPGTTNTNTSNNSNTNSQASQSNEDVDSLLKQLYALPLSQRHVESDGLIFDPAQITS  300

Query: 300  RTPNGVAIPHGDHYHFIPYSKLSALEEKIARRVPISGTGSTVSTNAKPNE  349
            RT  GVA+PHG+HYHFIPY ++S LEE+IAR +P+    +   +++P +
Sbjct: 301  RTARGVAVPHGNHYHFIPYEQMSELEERIARIIPLRYRSNHWVPDSRPEQ  350


  Score = 73.2 bits (178), Expect = 2e-11
  Identities = 73/288 (25%), Positives = 107/288 (36%), Gaps = 63/288 (21%)

Query: 280  SAQRYSESDGLVFDPAKIISRTPNGVAIPHGDHYHFIPYSKLSALEEKIARR--------  331
            S  RY+  DG +F+ + II  T +   +PHGDHYH+IP ++LSA E   A
Sbjct: 160  SQGRYTTDDGYIFNASDIIEDTGDAYIVPHGDHYHYIPKNELSASELAAAEAFLSGRGNL  219

Query: 332  -------------------VP-ISGTGST---VSTNAKPNEVVXXXXXXXXXXXXXXXXXK  368
                               VP +S G+T    S N+  N
Sbjct: 220  SNSRTYRLQNSDNTPRTNWVPSVSNPGTTNTNTSNNSNTNSQASQSNEDVDSLLKQLYAL  279

Query: 369  ELSSA---SDGYIFNPKDIVEETATAYIVRHGDHFHYIP-------------------KS  406
            LS      SDG IF+P  I    TA    V HG+H+H+IP                  +S
Sbjct: 280  PLSQRHVESDGLIFDPAQITSRTARGVAVPHGNHYHFIPYEQMSELEERIARIIPLRYRS  339

Query: 407  NQIGQPTLPNNSLATPSPS--LPINPGISHEKHEEDGYGFDANRIIAEDESGFIMSHGNH  464
            N    + P    PSPS    NP +      ++    +A R + +   G++
Sbjct: 340  NHWVPDSRPEQPSPQPSPSPQPAPNPQPAPSNPIDEKLVKEAVRKVGD---GYVFEENGV  396

Query: 465  NHYFFKKDLTEEQIKAAQKHL-EEVKTSHNGLDSLSSHEQDYPGNAKE  511
            + Y   KDL+ E        L ++   SH      L + + D P + +E
Sbjct: 397  SRYIPAKDLSAETAAGIDSKLAKQESLSH----KLGTKKTDLPSSDRE  440


  Score = 67.0 bits (162), Expect = 2e-09
  Identities = 48/190 (25%), Positives = 80/190 (41%), Gaps = 30/190 (15%)

Query: 138  LKDAAHADNVRTKDEIN------RQKQEHVKDNEKV---NSNVAVARSQGRYTTNDGYVF  188
            LKD + +D V+ +++I      R +  K N ++   +  + VA+ G+YTT DGY+F
Sbjct: 476  LKDVS-SDKVKLVEDILAFLAPIRHPERLGKPNSQITYTDDEIQVAKLAGKYTTEDGYIF  534

Query: 189  NPADIIEDTGNAYIVPHGGHYHYIPXXXXXXXXXXXXXXXXXXXXXNMQPSQLSYSSTASDN  248
            +P DI D G+AY+ PH  H H+I                    +  P    ST  +
Sbjct: 535  DPRDITSDEGDAYVTPHMTHSHWIKKDSLSEAERAAAQAYAKEKGLTP-----PSTDHQD  589

Query: 249  NTQSVAKGSTSKPANKSENLQSLLKELYDSPSAQRYSESDGLVFDPAKIISRTPNGVAIP  308
```

```
                  +   +  AKG+                    +  +Y+   A +    D + ++     +      + IP
Sbjct: 590 SGNTEAKGA---------------EAIYNRVKAAKKVPLDRMPYNLQYTVEVKNGSLIIP 634

Query: 309 HGDHYHFIPY 318
           H DHYH I +
Sbjct: 635 HYDHYHNIKF 644


  Score = 53.1 bits (126), Expect = 3e-05
  Identities = 57/315 (18%), Positives = 108/315 (34%), Gaps = 94/315 (29%)

Query: 297 IISRTPNGVAIPHGDHYHF----IPYSKLSALEEKIARRVPISGTGSTVSTNAKPNEVVX 352
           +I  T G    HGDHYH+    +PY + +
Sbjct: 51  VIKITDQGYVTSHGDHYHYYNGKVPYDAIFS----------------------------- 81

Query: 353 XXXXXXXXXXXXXXXKELSSASDGYIFNPKDIVEETATAYIVR------------HGDH 399
                          +EL      Y    +DIV E     Y+++           H D+
Sbjct: 82  --------------EELLMKDPNYKLKDEDIVNEVKGGYVIKVDGKYYVYLKDAAHADN 126

Query: 400 FHYIPKSNQIGQPTLPNNSLATPSPSLPINPGISHEKHE-EDGYGFDANRIIAEDESGFI 458
           + N+ Q   +    TP    +    S  ++ +DGY F+A+ II +    +I
Sbjct: 127 VRTKEEINRQKQEHSQHREGGTPRNDGAVALARSQGRYTTDDGYIFNASDIIEDTGDAYI 186

Query: 459 MSHGNHNHYFFKKDLTEEQIKAAQKHLE------------------------EVKTSH 492
           + HG+H HY  K +L+   ++ AA+  L                         V
Sbjct: 187 VPHGDHYHYIPKNELSASELAAAEAFLSGRGNLSNSRTYRLQNSDNTPRTNWVPSVSNPG 246

Query: 493 NGLDSLSSHEQDYPGNAKEMKDLDKKIEEKIAGIMKQYGVKRESI------VVNKEKNAI 546
           + S++        ++  +D+D +++  A + Q V+ + +      + ++    +
Sbjct: 247 TTNTNTSNNSNTNSQASQSNEDVDSLLKQLYALPLSQRHVESDGLIFDPAQITSRTARGV 306

Query: 547 IYPHGDHHHADPIDE 561
           PHG+H+H  P ++
Sbjct: 307 AVPHGNHYHFIPYEQ 321



  Score = 43.9 bits (102), Expect = 0.016
  Identities = 62/286 (21%), Positives = 95/286 (32%), Gaps = 43/286 (15%)

Query: 372 SASDGYIFNPKDIVEETATAYIVRHGDHFHYIPKSN------QIGQPTLPNNSLATPSPS 425
           +  DGYIF+P+DI  +   AY+  H  H H+I K +      Q      L   PS
Sbjct: 527 TTEDGYIFDPRDITSDEGDAYVTPHMTHSHWIKKDSLSEAERAAAQAYAKEKGLTPPSTD 586

Query: 426 LPINPGISHEKHEE-------------DGYGFDANRIIAEDESGFIMSHGNHNH----- 466
           + G + K E              D  ++   +     I+ H +H H
Sbjct: 587 HQ-DSGNTEAKGAEAIYNRVKAAKKVPLDRMPYNLQYTVEVKNGSLIIPHYDHYHNIKFE 645

Query: 467 ------YFFKKDLTEEQIKAAQK----HLEEVKTSHNGLDSLSSHEQDYPGNAKEMKDLD 516
                 Y   K   T E +A K     H  E   S NG  + S H +    +      D D
Sbjct: 646 WFDEGLYEAPKGYTLEDLLATVKYYVEHPNERPHSDNGFGNASDHVRKNKADQDSKPDED 705

Query: 517 KKIEEKIAGIMKQYGVKRESIVVNKEKNAIIYPHGDHHHA-----DPIDEHKPVGIGHSH 571
           K +E     + K    +N  + +  P D      D  DE +   + HS
Sbjct: 706 KGHDEVSEPTHPESDEKENHAGLNPSADNLYKPSTDTEETEEEAEDTTDEAEIPQVEHSV 765

Query: 572 SNYELFKPEEGVAKKEGNKVYTG--EELTNVVNLLKNSTFNNQNFT 615
           N ++    E  + K     +     E LT + + L   T +N   +
Sbjct: 766 INAKIADAEALLEKVTDPSIRQNAMETLTGLKSSLLLGTKDNNTIS 811
```

```
Score = 39.7 bits (91), Expect = 0.30
Identities = 34/138 (24%), Positives = 51/138 (36%), Gaps = 19/138 (13%)

Query: 274  ELYDSPSAQRYSESDGLVFDPAKIISRTPNGVAIPHGDHYHFIPYSKLSALEEKIARRV- 332
            E+  +  A +Y+  DG +FDP  I S  +    PH  H H+I   LS  E   A+
Sbjct: 516  EIQVAKLAGKYTTEDGYIFDPRDITSDEGDAYVTPHMTHSHWIKKDSLSEAERAAAQAYA 575

Query: 333  -------PISGTGSTVSTNAKPNEVVXXXXXXXXXXXXXXXXXXXKELSSASDGYIFNPKDIV 385
                   P +   + +T AK  E +                         D  +N + V
Sbjct: 576  KEKGLTPPSTDHQDSGNTEAKGAEAIYNRVKAAK----------KVPLDRMPYNLQYTV 624

Query: 386  EETATAYIVRHGDHFHYI 403
            E    + I+ H DH+H I
Sbjct: 625  EVKNGSLIIPHYDHYHNI 642
```

tr <u>Q6WNP5</u> **Surface protein BVH-11-2 (Fragment) [bvh-11-2] [Streptococcus** 830 AA
         pneumoniae]

<u>align</u>

```
Score =  415 bits (1066), Expect = e-114
Identities = 206/350 (58%), Positives = 257/350 (72%), Gaps = 21/350 (6%)

Query: 21   CAYALNQHRSQENKDNNRVSYVDGSQSSQKSENLTPDQVSQKEGIQAEQIVIKITDQGYV 80
            C+Y L ++++  K+NNRVSY+DG Q++QK+ENLTPD+VS++EGI AEQIVIKITDQGYV
Sbjct: 1    CSYELGLYQARTVKENNRVSYIDGKQATQKTENLTPDEVSKREGINAEQIVIKITDQGYV 60

Query: 81   TSHGDHYHYYNGKVPYDALFSEELLMKDPNYQLKDADIVNEVKGGYIIKVDGKYYVYLKD 140
            TSHGDHYHYYNGKVPYDA+FSEELLMKDPNY+LKD DIVNEVKGGY+IKVDGKYYVYLKD
Sbjct: 61   TSHGDHYHYYNGKVPYDAIFSEELLMKDPNYKLKDEDIVNEVKGGYVIKVDGKYYVYLKD 120

Query: 141  AAHADNVRTKDEINRQKQEHVKDNE----KVNSNVAVARSQGRYTTNDGYVFNPADIIED 196
            AAHADNVRTK+EINRQKQEH +  E   + +  VA+ARSQGRYTT+DGY+FN +DIIED
Sbjct: 121  AAHADNVRTKEEINRQKQEHSQHREGGTPRNDGAVALARSQGRYTTDDGYIFNASDIIED 180

Query: 197  TGNAYIVPHGGHYHYIPXXXXXXXXXXXXXXXXXXXXXXNMQPSQLSYSSTASDN-------- 248
            TG+AYIVPHG HYHYIP                       +  +Y   SDN
Sbjct: 181  TGDAYIVPHGDHYHYIPKNELSASELAAAEAFLSGRGNLSNSRTYRLQNSDNTPRTNWVP 240

Query: 249  ---------NTQSVAKGSTSKPANKS-ENLQSLLKELYDSPSAQRYSESDGLVFDPAKIIS 299
                     NT +    +T+ A++S E++ SLLK+LY P +QR+ ESDGL+FDPA+I S
Sbjct: 241  SVSNPGTTNTNTSNNSNTNSQASQSNEDVDSLLKQLYALPLSQRHVESDGLIFDPAQITS 300

Query: 300  RTPNGVAIPHGDHYHFIPYSKLSALEEKIARRVPISGTGSTVSTNAKPNE 349
            RT  GVA+PHG+HYHFIPY ++S LEE+IAR +P+     +    +++P +
Sbjct: 301  RTARGVAVPHGNHYHFIPYEQMSELEERIARIIPLRYRSNHWVPDSRPEQ 350
```

```
Score = 73.2 bits (178), Expect = 2e-11
Identities = 73/288 (25%), Positives = 107/288 (36%), Gaps = 63/288 (21%)

Query: 280  SAQRYSESDGLVFDPAKIISRTPNGVAIPHGDHYHFIPYSKLSALEEKIARR-------- 331
            S  RY+  DG +F+  + II  T +   +PHGDHYH+IP ++LSA E   A
Sbjct: 160  SQGRYTTDDGYIFNASDIIEDTGDAYIVPHGDHYHYIPKNELSASELAAAEAFLSGRGNL 219
```

```
Query: 332 ------------------VP-ISGTGST---VSTNAKPNEVVXXXXXXXXXXXXXXXXXK 368
                             VP +S G+T    S N+  N
Sbjct: 220 SNSRTYRLQNSDNTPRTNWVPSVSNPGTTNTNTSNNSNTNSQASQSNEDVDSLLKQLYAL 279

Query: 369 ELSSA---SDGYIFNPKDIVEETATAYIVRHGDHFHYIP------------------KS 406
             LS      SDG IF+P  I   TA   V HG+H+H+IP                  +S
Sbjct: 280 PLSQRHVESDGLIFDPAQITSRTARGVAVPHGNHYHFIPYEQMSELEERIARIIPLRYRS 339

Query: 407 NQIGQPTLPNNSLATPSPS--LPINPGISHEKHEEDGYGFDANRIIAEDESGFIMSHGNH 464
             N    + P      PSPS    NP +    ++    +A R + +    G++
Sbjct: 340 NHWVPDSRPEQPSPQPSPSPQPAPNPQPAPSNPIDEKLVKEAVRKVGD---GYVFEENGV 396

Query: 465 NHYFFKKDLTEEQIKAAQKHL-EEVKTSHNGLDSLSSHEQDYPGNAKE 511
             + Y  KDL+ E        L ++   SH      L + + D P + +E
Sbjct: 397 SRYIPAKDLSAETAAGIDSKLAKQESLSH----KLGTKKTDLPSSDRE 440


 Score = 67.0 bits (162), Expect = 2e-09
 Identities = 48/190 (25%), Positives = 80/190 (41%), Gaps = 30/190 (15%)

Query: 138 LKDAAHADNVRTKDEIN------RQKQEHVKDNEKV---NSNVAVARSQGRYTTNDGYVF 188
             LKD + +D V+ ++I      R +   K N ++   +   VA+  G+YTT DGY+F
Sbjct: 476 LKDVS-SDKVKLVEDILAFLAPIRHPERLGKPNSQITYTDDEIQVAKLAGKYTTEDGYIF 534

Query: 189 NPADIIEDTGNAYIVPHGGHYHYIPXXXXXXXXXXXXXXXXXXXXXXNMQPSQLSYSSTASDN 248
             +P DI  D G+AY+ PH  H H+I                      + P      ST   +
Sbjct: 535 DPRDITSDEGDAYVTPHMTHSHWIKKDSLSEAERAAAQAYAKEKGLTP-----PSTDHQD 589

Query: 249 NTQSVAKGSTSKPANKSENLQSLLKELYDSPSAQRYSESDGLVFDPAKIISRTPNGVAIP 308
             + + AKG+              + +Y+   A +    D + ++    +        + IP
Sbjct: 590 SGNTEAKGA--------------EAIYNRVKAAKKVPLDRMPYNLQYTVEVKNGSLIIP 634

Query: 309 HGDHYHFIPY 318
             H DHYH I +
Sbjct: 635 HYDHYHNIKF 644


 Score = 55.5 bits (132), Expect = 5e-06
 Identities = 69/373 (18%), Positives = 129/373 (34%), Gaps = 108/373 (28%)

Query: 239 LSYSSTASDNNTQSVAKGSTSKPANKSENLQSLLKELYDSPSAQRYSESDGLVFDPAKII 298
             L + T +NN S  G +    K+ENL          +    S+ +G+ +    +I
Sbjct: 7   LYQARTVKENNRVSYIDGKQA--TQKTENL----------TPDEVSKREGINAEQI-VI 52

Query: 299 SRTPNGVAIPHGDHYHF----IPYSKLSALEEKIARRVPISGTGSTVSTNAKPNEVVXXX 354
              T  G    HGDHYH+     +PY + +
Sbjct: 53  KITDQGYVTSHGDHYHYYNGKVPYDAIFS------------------------------ 81

Query: 355 XXXXXXXXXXXXXXXKELSSASDGYIFNPKDIVEETATAYIVR------------HGDHFH 401
                             +EL    Y    +DIV E    Y+++            H D+
Sbjct: 82  ------------EELLMKDPNYKLKDEDIVNEVKGGYVIKVDGKYYVYLKDAAHADNVR 128

Query: 402 YIPKSNQIGQPTLPNNSLATPSPSLPINPGISHEKHE-EDGYGFDANRIIAEDESGFIMS 460
              + N+  Q  +      TP     S  ++  +DGY F+A+ II +     +1+
Sbjct: 129 TKEEINRQKQEHSQHREGGTPRNDGAVALARSQGRYTTDDGYIFNASDIIEDTGDAYIVP 188

Query: 461 HGNHNHYFFKKDLTEEQIKAAQKHLE------------------------EVKTSHNG 494
```

```
                       HG+H HY   K +L+   ++ AA+   L                                    V
Sbjct: 189  HGDHYHYIPKNELSASELAAAEAFLSGRGNLSNSRTYRLQNSDNTPRTNWVPSVSNPGTT 248

Query: 495  LDSLSSHEQDYPGNAKEMKDLDKKIEEKIAGIMKQYGVKRESI------VVNKEKNAIIY 548
              + S++         ++  +D+D +++ A  + Q V++ +        + ++      +
Sbjct: 249  NTNTSNNSNTNSQASQSNEDVDSLLKQLYALPLSQRHVESDGLIFDPAQITSRTARGVAV 308

Query: 549  PHGDHHHADPIDE 561
            PHG+H+H  P ++
Sbjct: 309  PHGNHYHFIPYEQ 321


  Score = 43.9 bits (102), Expect = 0.016
  Identities = 62/286 (21%), Positives = 95/286 (32%), Gaps = 43/286 (15%)

Query: 372  SASDGYIFNPKDIVEETATAYIVRHGDHFHYIPKSN------QIGQPTLPNNSLATPSPS 425
             +  DGYIF+P+DI  +    AY+ H  H H+I K +         Q      L  PS
Sbjct: 527  TTEDGYIFDPRDITSDEGDAYVTPHMTHSHWIKKDSLSEAERAAAQAYAKEKGLTPPSTD 586

Query: 426  LPINPGISHEKHEE-------------DGYGFDANRIIAEDESGFIMSHGNHNH----- 466
              + G +  K E             D  ++       +      I+ H +H H
Sbjct: 587  HQ-DSGNTEAKGAEAIYNRVKAAKKVPLDRMPYNLQYTVEVKNGSLIIPHYDHYHNIKFE 645

Query: 467  ------YFFKKDLTEEQIKAAQK----HLEEVKTSHNGLDSLSSHEQDYPGNAKEMKDLD 516
                  Y   K  T E+ A K       H E    S NG  + S H +      +   D D
Sbjct: 646  WFDEGLYEAPKGYTLEDLLATVKYYVEHPNERPHSDNGFGNASDHVRKNKADQDSKPDED 705

Query: 517  KKIEEKIAGIMKQYGVKRESIVVNKEKNAIIYPHGDHHHA-----DPIDEHKPVGIGHSH 571
            K +E        +    K     +N  + ++  P   D        D  DE +  + HS
Sbjct: 706  KGHDEVSEPTHPESDEKENHAGLNPSADNLYKPSTDTEETEEEAEDTTDEAEIPQVEHSV 765

Query: 572  SNYELFKPEEGVAKKEGNKVYTG--EELTNVVNLLKNSTFNNQNFT 615
             N ++    E + K      +    E LT ++ + L   T +N  +
Sbjct: 766  INAKIADAEALLEKVTDPSIRQNAMETLTGLKSSLLLGTKDNNTIS 811


  Score = 39.7 bits (91), Expect = 0.30
  Identities = 34/138 (24%), Positives = 51/138 (36%), Gaps = 19/138 (13%)

Query: 274  ELYDSPSAQRYSESDGLVFDPAKIISRTPNGVAIPHGDHYHFIPYSKLSALEEKIARRV- 332
            E+ +  A +Y+  DG +FDP  I S   +    PH  H  H+I    LS  E    A+
Sbjct: 516  EIQVAKLAGKYTTEDGYIFDPRDITSDEGDAYVTPHMTHSHWIKKDSLSEAERAAAQAYA 575

Query: 333  -------PISGTGSTVSTNAKPNEVVXXXXXXXXXXXXXXXXXXXKELSSASDGYIFNPKDIV 385
                   P +      + +T AK  E +                           D  +N + V
Sbjct: 576  KEKGLTPPSTDHQDSGNTEAKGAEAIYNRVKAAK-----------KVPLDRMPYNLQYTV 624

Query: 386  EETATAYIVRHGDHFHYI 403
            E    + I+ H DH+H I
Sbjct: 625  EVKNGSLIIPHYDHYHNI 642



tr  Q6WNP9  Surface protein BVH-11-2 (Fragment) [Streptococcus pneumoniae]  819 AA

                                                                    align
```

```
 Score =  410 bits (1054), Expect = e-113
 Identities = 204/337 (60%), Positives = 255/337 (75%), Gaps = 10/337 (2%)

Query: 21   CAYALNQHRS-QENKDNNRVSYVDGSQSSQKSENLTPDQVSQKEGIQAEQIVIKITDQGY 79
            C+Y L +H++ Q  K++NRVSY+DG Q+ QK+ENLTPD+VS++EGI AEQIVIKITDQGY
Sbjct: 1    CSYELGRHQAGQVKKESNRVSYIDGDQAGQKAENLTPDEVSKREGINAEQIVIKITDQGY 60

Query: 80   VTSHGDHYHYYNGKVPYDALFSEELLMKDPNYQLKDADIVNEVKGGYIIKVDGKYYVYLK 139
            VTSHGDHYHYYNGKVPYDA+ SEELLMKDPNYQLKD+DIVNE+KGGY+IKVDGKYYVYLK
Sbjct: 61   VTSHGDHYHYYNGKVPYDAIISEELLMKDPNYQLKDSDIVNEIKGGYVIKVDGKYYVYLK 120

Query: 140  DAAHADNVRTKDEINRQKQEHVKD-NEKVNSNVAVARSQGRYTTNDGYVFNPADIIEDTG 198
            DAAHADN+RTK+EI RQKQE   + N + ++ VA AR+QGRYTT+DGY+FN +DIIEDTG
Sbjct: 121  DAAHADNIRTKEEIKRQKQERSHNHNSRADNAVAAARAQGRYTTDDGYIFNASDIIEDTG 180

Query: 199  NAYIVPHGGHYHYIPXXXXXXXXXXXXXXXXXXXXXXNMQPSQLSYSSTASDNNTQ---SVAK 255
            +AYIVPHG H+HYIP                       Q S+ S SS+ + N  Q     S
Sbjct: 181  DAYIVPHGNHFHYIP--KSDLSASELAAAQAYWNGKQGSRPSSSSSHNANPAQPRLSENH 238

Query: 256  GSTSKPA---NKSENLQSLLKELYDSPSAQRYSESDGLVFDPAKIISRTPNGVAIPHGDH 312
               T  P    N+ EN+ SLL+ELY  P ++R+ ESDGL+FDPA+I SRT  GVA+PHG+H
Sbjct: 239  NLTVTPTYHQNQGENISSLLRELYAKPLSERHVESDGLIFDPAQITSRTARGVAVPHGNH 298

Query: 313  YHFIPYSKLSALEEKIARRVPISGTGSTVSTNAKPNE 349
            YHFIPYS++S LEE+IAR +P+      +   +++P +
Sbjct: 299  YHFIPYSQMSELEERIARIIPLRYRSNHWVPDSRPEQ 335


 Score = 77.8 bits (190), Expect = 1e-12
 Identities = 86/385 (22%), Positives = 135/385 (34%), Gaps = 73/385 (18%)

Query: 193  IIEDTGNAYIVPHGGHYHYI------------------PXXXXXXXXXXXXXXXXXXXXXNM 234
            +I+ T  Y+  HG HYHY                    P                      +
Sbjct: 52   VIKITDQGYVTSHGDHYHYYNGKVPYDAIISEELLMKDPNYQLKDSDIVNEIKGGYVIKV 111

Query: 235  QPSQLSYSSTASDNNTQSVAKGSTSKPANKSENLQSLLKELYDSPSAQ-RYSESDGLVFD 293
                Y  A+ +    +   +S N  S     + AQ RY+ DG +F+
Sbjct: 112  DGKYYVYLKDAAHADNIRTKEEIKRQKQERSHNHNSRADNAVAAARAQGRYTTDDGYIFN 171

Query: 294  PAKIISRTPNGVAIPHGDHYHFIPYSKLSALEEKIARRVPISGTGSTVST----NAKP-- 347
            + II T +   +PHG+H+H+IP S LSA E   A+       GS  S+    NA P
Sbjct: 172  ASDIIEDTGDAYIVPHGNHFHYIPKSDLSASELAAAQAYWNGKQGSRPSSSSSHNANPAQ 231

Query: 348  -------NEVVXXXXXXXXXXXXXXXXXKELSSA--------SDGYIFNPKDIVEETATAY 392
                   N  V                +EL +         SDG IF+P I   TA
Sbjct: 232  PRLSENHNLTVTPTYHQNQGENISSLLRELYAKPLSERHVESDGLIFDPAQITSRTARGV 291

Query: 393  IVRHGDHFHYIPKS------------------------NQIGQPTLPNNSLATPSPSLP 427
            V HG+H+H+IP S                         ++  QP+ +      +PSP
Sbjct: 292  AVPHGNHYHFIPYSQMSELEERIARIIPLRYRSNHWVPDSRPEQPSPQSTPEPSPSPQSA 351

Query: 428  INPGISHEKHEEDGYGFDANRIIAEDESGFIMSHGNHNHYFFKKDLTEEQIKAAQKHL-E 486
            NP +      ++  +A R + + G++      + Y   K+L+ E      L +
Sbjct: 352  PNPQPAPSNPIDEKLVKEAVRKVGD---GYVFEKNGVSRYIPAKNLSAETAAGIDSKLAK 408

Query: 487  EVKTSHNGLDSLSSHEQDYPGNAKE 511
            +  SH    L + + D P + +E
Sbjct: 409  QESLSH----KLGAKKTDLPSSDRE 429
```

```
Score = 73.6 bits (179), Expect = 2e-11
Identities = 110/550 (20%), Positives = 205/550 (37%), Gaps = 79/550 (14%)

Query:  29 RSQENKDNNRVSYVDGSQSSQKSENLTPDQVSQKEG-IQAEQIVIKITDQGYVTSHGDHY  87
           R ++ + +N  S  D + ++ +++     + +  +G I     +I+ T   Y+  HG+H+
Sbjct: 136 RQKQERSHNHNSRADNAVAAARAQG----RYTTDDGYIFNASDIIEDTGDAYIVPHGNHF 191

Query:  88 HYYNGKVPYDALFSEELLMKDPNYQLKDADIVNEVKGGYIIKVDGKYYVYLKDAAHADNV 147
           HY    +P   L + EL      + K     +            +         + N+
Sbjct: 192 HY----IPKSDLSASELAAAQAYWNGKQGSRPSSSSSHNANPAQPRL-------SENHNL 240

Query: 148 RTKDEINRQKQEHVKDNEKVNSNVAVARSQGRYTTNDGYVFNPADIIEDTGNAYIVPHGG 207
              ++ + E++ + +   A  S+ R+ +DG +F+PA I   T      VPHG
Sbjct: 241 TVTPTYHQNQGENI--SSLLRELYAKPLSE-RHVESDGLIFDPAQITSRTARGVAVPHGN 297

Query: 208 HYHYIPXXXXXXXXXXXXXXXXXXX-------NMQPSQLSYSSTASDNNTQSVAKGSTSK 260
           HYH+IP                          + +P Q S  ST  + +    A
Sbjct: 298 HYHFIPYSQMSELEERIARIIPLRYRSNHWVPDSRPEQPSPQSTPEPSPSPQSAPNPQPA 357

Query: 261 PANKSENLQSLLKELYDSPSAQRYSESDGLVFDPAKIISRTPNGVAIPHGDHYHFIPYSK 320
           P+N  + + L+KE       DG VF+         NGV+       +IP
Sbjct: 358 PSNPID--EKLVKEAVRKVG------DGYVFEK--------NGVS-------RYIPAKN 393

Query: 321 LSA-----LEEKIARRVPISGTGSTVSTNAKP------NEVVXXXXXXXXXXXXXXXXXKE 369
           LSA      ++ K+A++  +S      T+        N+                   +
Sbjct: 394 LSAETAAGIDSKLAKQESLSHKLGAKKTDLPSSDREFYNKAYDLLARIHQDLLDNKGRQV 453

Query: 370 LSSASDGYIFNPKDIVEETATAYIVRHGDHFHYIPKSNQIGQPTLPNNSLATPSPSLPIN 429
              A D +   KD+ +    D  ++      +   PN +     + +
Sbjct: 454 DFEALDNLLERLKDVPSDKVKLV----DDILAFLAPIRHPERLGKPNAQITYTDDEIQVA 509

Query: 430 PGISHEKHEEDGYGFDANRIIAEDESGFIMSHGNHNHYFFKKDLTEEQIKAAQKHLEEVK 489
           ++ +     EDGY FD   I +++   ++   H  H+H+   K  L+E +  AAQ + +E
Sbjct: 510 K-LAGKYTTEDGYIFDPRDITSDEGDAYVTPHMTHSHWIKKDSLSEAERAAAQAYAKE-- 566

Query: 490 TSHNGLDSLSSHEQDYPGN--AKEMKDLDKKIE--EKIAGIMKQYGVKRESIVVNKEKNA 545
             GL   S+   QD   GN  AK + +  +++    +K+    Y ++        V +  +
Sbjct: 567 ---KGLTPPSTDHQD-SGNTEAKGAEAIYNRVKAAKKVPLDRMPYNLQ---YTVEVKNGS 619

Query: 546 IIYPHGDHHH 555
           +I PH DH+H
Sbjct: 620 LIIPHYDHYH 629
```

```
Score = 67.4 bits (163), Expect = 1e-09
Identities = 49/190 (25%), Positives = 79/190 (40%), Gaps = 30/190 (15%)

Query: 138 LKDAAHADNVRTKDEIN------RQKQEHVKDNEKV---NSNVAVARSQGRYTTNDGYVF 188
           LKD   +D V+  D+I       R +   K N ++  +   + VA+  G+YTT DGY+F
Sbjct: 465 LKDVP-SDKVKLVDDILAFLAPIRHPERLGKPNAQITYTDDEIQVAKLAGKYTTEDGYIF 523

Query: 189 NPADIIEDTGNAYIVPHGGHYHYIPXXXXXXXXXXXXXXXXXXXXXNMQPSQLSYSSTASDN 248
           +P DI  D G+AY+ PH  H H+I                       + P     ST  +
Sbjct: 524 DPRDITSDEGDAYVTPHMTHSHWIKKDSLSEAERAAAQAYAKEKGLTP-----PSTDHQD 578

Query: 249 NTQSVAKGSTSKPANKSENLQSLLKELYDSPSAQRYSESDGLVFDPAKIISRTPNGVAIP 308
```

```
                  +   + AKG+                  +  +Y+    A  +    D + ++     +        + IP
Sbjct: 579 SGNTEAKGA--------------EAIYNRVKAAKKVPLDRMPYNLQYTVEVKNGSLIIP 623

Query: 309 HGDHYHFIPY 318
           H DHYH I +
Sbjct: 624 HYDHYHNIKF 633
```

```
 Score = 53.5 bits (127), Expect = 2e-05
 Identities = 58/302 (19%), Positives = 109/302 (35%), Gaps = 84/302 (27%)

Query: 297 IISRTPNGVAIPHGDHYHF----IPYSKLSALEEKIARRVPISGTGSTVSTNAKPNEVVX 352
           +I  T  G    HGDHYH+     +PY   + +
Sbjct: 52  VIKITDQGYVTSHGDHYHYYNGKVPYDAIIS--------------------------- 82

Query: 353 XXXXXXXXXXXXXXXXKELSSASDGYIFNPKDIVEETATAYIVR-HGDHFHYIPKS----- 406
                           +EL       Y     DIV E   Y+++  G ++ Y+   +
Sbjct: 83  --------------EELLMKDPNYQLKDSDIVNEIKGGYVIKVDGKYYVYLKDAAHADN 127

Query: 407 --------NQIGQPTLPNNSLATPSPSLPINPGISHEKHEEDGYGFDANRIIAEDESGFI 458
                   Q  + +NS A + +      G     +DGY F+A+ II +    +I
Sbjct: 128 IRTKEEIKRQKQERSHNHNSRADNAVAAARAQG---RYTTDDGYIFNASDIIEDTGDAYI 184

Query: 459 MSHGNHNHYFFKKDLTEEQIKAAQKHLE-------EVKTSHNGLDSLSSHEQDY------ 505
           + HGNH HY  K DL+  ++ AAQ +        +SHN    +   +++
Sbjct: 185 VPHGNHFHYIPKSDLSASELAAAQAYWNGKQGSRPSSSSSHNANPAQPRLSENHNLTVTP 244

Query: 506 PGNAKEMKDLDKKIEEKIAGIMKQYGVKRESI------VVNKEKNAIIYPHGDHHHADPI 559
             +  + +++     + E  A  + + V+ + +       + ++     + PHG+H+H  P
Sbjct: 245 TYHQNQGENISSLLRELYAKPLSERHVESDGLIFDPAQITSRTARGVAVPHGNHYHFIPY 304

Query: 560 DE 561
            +
Sbjct: 305 SQ 306
```

```
 Score = 43.1 bits (100), Expect = 0.027
 Identities = 44/185 (23%), Positives = 64/185 (33%), Gaps = 36/185 (19%)

Query: 372 SASDGYIFNPKDIVEETATAYIVRHGDHFHYIPKSN------QIGQPTLPNNSLATPSPS 425
           +  DGYIF+P+DI  +   AY+ H H H+I K +        Q      L  PS
Sbjct: 516 TTEDGYIFDPRDITSDEGDAYVTPHMTHSHWIKKDSLSEAERAAAQAYAKEKGLTPPSTD 575

Query: 426 LPINPGISHEKHEE--------------DGYGFDANRIIAEDESGFIMSHGNHNH----- 466
           + G + K E               D  ++   +        I+ H +H H
Sbjct: 576 HQ-DSGNTEAKGAEAIYNRVKAAKKVPLDRMPYNLQYTVEVKNGSLIIPHYDHYHNIKFE 634

Query: 467 ------YFFKKDLTEEQIKAAQK----HLEEVKTSHNGLDSLSSHEQDYPGNAKEMKDLD 516
                 Y  K  TE +A K    H E  S NG  + S H +   +     D D
Sbjct: 635 WFDEGLYEAPKGYTLEDLLATVKYYVEHPNERPHSDNGFGNASDHVRKNKADQDSKPDED 694

Query: 517 KKIEE 521
           K+ +E
Sbjct: 695 KEHDE 699
```

```
 Score = 39.7 bits (91), Expect = 0.30
 Identities = 34/138 (24%), Positives = 51/138 (36%), Gaps = 19/138 (13%)
```

```
Query:  274  ELYDSPSAQRYSESDGLVFDPAKIISRTPNGVAIPHGDHYHFIPYSKLSALEEKIARRV-  332
             E+  +   A +Y+   DG +FDP   I S   +     PH  H H+I    LS  E    A+
Sbjct:  505  EIQVAKLAGKYTTEDGYIFDPRDITSDEGDAYVTPHMTHSHWIKKDSLSEAERAAAQAYA  564

Query:  333  -------PISGTGSTVSTNAKPNEVVXXXXXXXXXXXXXXXXXXKELSSASDGYIFNPKDIV  385
                     P +    + +T AK  E +                     D  +N +  V
Sbjct:  565  KEKGLTPPSTDHQDSGNTEAKGAEAIYNRVKAAK----------KVPLDRMPYNLQYTV  613

Query:  386  EETATAYIVRHGDHFHYI  403
             E    + I+ H DH+H I
Sbjct:  614  EVKNGSLIIPHYDHYHNI  631
```

tr   <u>Q6T304</u>   **Surface protein BVH-11 (Fragment) [bvh11]**   807 AA
                        [Streptococcus   <u>align</u>
                        pneumoniae]

```
 Score =  410 bits (1054), Expect = e-113
 Identities = 204/337 (60%), Positives = 255/337 (75%), Gaps = 10/337 (2%)

Query:  21   CAYALNQHRS-QENKDNNRVSYVDGSQSSQKSENLTPDQVSQKEGIQAEQIVIKITDQGY  79
             C+Y L +H++ Q  K++NRVSY+DG Q+ QK+ENLTPD+VS++EGI AEQIVIKITDQGY
Sbjct:  1    CSYELGRHQAGQVKKESNRVSYIDGDQAGQKAENLTPDEVSKREGINAEQIVIKITDQGY  60

Query:  80   VTSHGDHYHYYNGKVPYDALFSEELLMKDPNYQLKDADIVNEVKGGYIIKVDGKYYVYLK  139
             VTSHGDHYHYYNGKVPYDA+  SEELLMKDPNYQLKD+DIVNE+KGGY+IKVDGKYYVYLK
Sbjct:  61   VTSHGDHYHYYNGKVPYDAIISEELLMKDPNYQLKDSDIVNEIKGGYVIKVDGKYYVYLK  120

Query:  140  DAAHADNVRTKDEINRQKQEHVKD-NEKVNSNVAVARSQGRYTTNDGYVFNPADIIEDTG  198
             DAAHADN+RTK+EI RQKQE   + N + ++ VA AR+QGRYTT+DGY FN +DIIEDTG
Sbjct:  121  DAAHADNIRTKEEIKRQKQERSHNHNSRADNAVAAARAQGRYTTDDGYIFNASDIIEDTG  180

Query:  199  NAYIVPHGGHYHYIPXXXXXXXXXXXXXXXXXXXXXNMQPSQLSYSSTASDNNTQ---SVAK  255
             +AYIVPHG H+HYIP                     Q S+ S SS+ + N  Q    S
Sbjct:  181  DAYIVPHGNHFHYIP--KSDLSASELAAAQAYWNGKQGSRPSSSSSHNANPAQPRLSENH  238

Query:  256  GSTSKPA---NKSENLQSLLKELYDSPSAQRYSESDGLVFDPAKIISRTPNGVAIPHGDH  312
                T   P    N+ EN+ SLL+ELY  P ++R+ ESDGL+FDPA+I SRT GVA+PHG+H
Sbjct:  239  NLTVTPTYHQNQGENISSLLRELYAKPLSERHVESDGLIFDPAQITSRTARGVAVPHGNH  298

Query:  313  YHFIPYSKLSALEEKIARRVPISGTGSTVSTNAKPNE  349
             YHFIPYS++S LEE+IAR +P+     +   +++P +
Sbjct:  299  YHFIPYSQMSELEERIARIIPLRYRSNHWVPDSRPEQ  335


 Score = 76.6 bits (187), Expect = 2e-12
 Identities = 84/385 (21%), Positives = 133/385 (33%), Gaps = 73/385 (18%)

Query:  193  IIEDTGNAYIVPHGGHYHYI------------------PXXXXXXXXXXXXXXXXXXXXXNM  234
             +I+ T  Y+  HG HYHY                   P                   +
Sbjct:  52   VIKITDQGYVTSHGDHYHYYNGKVPYDAIISEELLMKDPNYQLKDSDIVNEIKGGYVIKV  111

Query:  235  QPSQLSYSSTASDNNTQSVAKGSTSKPANKSENLQSLLKELYDSPSAQ-RYSESDGLVFD  293
                    Y  A+  +     +     +S N  S       + AQ RY+  DG +F+
Sbjct:  112  DGKYYVYLKDAAHADNIRTKEEIKRQKQERSHNHNSRADNAVAAARAQGRYTTDDGYIFN  171
```

```
Query:  294  PAKIISRTPNGVAIPHGDHYHFIPYSKLSALEEKIARRVPISGTGSTVST----NAKP--  347
             +  II  T  +     +PHG+H+H+IP S LSA E     A+      GS   S+    NA P
Sbjct:  172  ASDIIEDTGDAYIVPHGNHFHYIPKSDLSASELAAAQAYWNGKQGSRPSSSSSHNANPAQ  231

Query:  348  -------NEVVXXXXXXXXXXXXXXXXXXKELSSA--------SDGYIFNPKDIVEETATAY  392
                    N  V                   +EL +       SDG IF+P  I    TA
Sbjct:  232  PRLSENHNLTVTPTYHQNQGENISSLLRELYAKPLSERHVESDGLIFDPAQITSRTARGV  291

Query:  393  IVRHGDHFHYIPKS-----------------------NQIGQPTLPNNSLATPSPSLP  427
             V HG+H+H+IP S                        ++  QP+  +     +PSP
Sbjct:  292  AVPHGNHYHFIPYSQMSELEERIARIIPLRYRSNHWVPDSRPEQPSPQSTPEPSPSPQSA  351

Query:  428  INPGISHEKHEEDGYGFDANRIIAEDESGFIMSHGNHNHYFFKKDLTEEQIKAAQKHL-E  486
             NP  +     ++      ++ +    G++      + Y  K+L+ E           L +
Sbjct:  352  PNPQPAPSNPIDEKL---VKEVVRKVGDGYVFEKNGVSRYIPAKNLSAETAAGIDSKLAK  408

Query:  487  EVKTSHNGLDSLSSHEQDYPGNAKE  511
             +    SH     L + + D P + +E
Sbjct:  409  QESLSH----KLGAKKTDLPSSDRE  429
```

```
Score = 75.1 bits (183), Expect = 7e-12
Identities = 108/546 (19%), Positives = 207/546 (37%), Gaps = 71/546 (13%)
```

```
Query:   29  RSQENKDNNRVSYVDGSQSSQKSENLTPDQVSQKEG-IQAEQIVIKITDQGYVTSHGDHY  87
             R ++ + +N  S  D + ++ +++      + +  +G I    +I+  T   Y+   HG+H+
Sbjct:  136  RQKQERSHNHNSRADNAVAAARAQG----RYTTDDGYIFNASDIIEDTGDAYIVPHGNHF  191

Query:   88  HYYNGKVPYDALFSEELLMKDPNYQLKDADIVNEVKGGYIIKVDGKYYVYLKDAAHADNV  147
             HY    +P   L + EL      + K    +              +        +    N+
Sbjct:  192  HY----IPKSDLSASELAAAQAYWNGKQGSRPSSSSSHNANPAQPRL-------SENHNL  240

Query:  148  RTKDEINRQKQEHVKDNEKVNSNVAVARSQGRYTTNDGYVFNPADIIEDTGNAYIVPHGG  207
             ++ + E+++ + +     A   S+ R+ +DG +F+PA I   T      VPHG
Sbjct:  241  TVTPTYHQNQGENI--SSLLRELYAKPLSE-RHVESDGLIFDPAQITSRTARGVAVPHGN  297

Query:  208  HYHYIPXXXXXXXXXXXXXXXXXXXXX-------NMQPSQLSYSSTASDNNTQSVAKGSTSK  260
             HYH+IP                           + +P Q S  ST  + +    A
Sbjct:  298  HYHFIPYSQMSELEERIARIIPLRYRSNHWVPDSRPEQPSPQSTPEPSPSPQSAPNPQPA  357

Query:  261  PANKSENLQSLLKELYDSPSAQRYSESDGLVFDPAKIISRTPNGVAIPHGDHYHFIPYSK  320
             P+N + + + L+KE+          DG VF+       NGV+       +IP
Sbjct:  358  PSNPID--EKLVKEVVRKVG-------DGYVFEK--------NGVS-------RYIPAKN  393

Query:  321  LSA-----LEEKIARRVPISGTGSTVSTN--AKPNEVVXXXXXXXXXXXXXXXXXXKELSSA  373
             LSA     ++ K+A++  +S       T+ +   E                          +
Sbjct:  394  LSAETAAGIDSKLAKQESLSHKLGAKKTDLPSSDREFYNKAYDLLARIHQDLLDNKGRQV  453

Query:  374  SDGYIFNPKDIVEETATAYIVRHGDHFHYIPKSNQIGQPTLPNNSLATPSPSLPINPGIS  433
             + N  + +E+ +  +    D  ++          +   PN  +    + +  ++
Sbjct:  454  DFEALDNLLERLEDVPSDKVKLVDDILAFLAPIRHPERLGKPNAQITYTDDEIQVAK-LA  512

Query:  434  HEKHEEDGYGFDANRIIAEDESGFIMSHGNHNHYFFKKDLTEEQIKAAQKHLEEVKTSHN  493
             +   EDGY FD   I +++    ++  H  H+H+  K  L+E +  AAQ + +E
Sbjct:  513  GKYTTEDGYIFDPRDITSDEGDAYVTPHMTHSHWIKKDSLSEAERAAAQAYAKE-----K  567

Query:  494  GLDSLSSHEQDYPGN--AKEMKDLDKKIE--EKIAGIMKQYGVKRESIVVNKEKNAIIYP  549
```

```
              GL    S+   QD   GN   AK  + +   +++    +K+       Y ++      V  +  ++I P
Sbjct: 568 GLTPPSTDHQD-SGNTEAKGAEAIYNRVKAAKKVPLDRMPYNLQ---YTVEVKNGSLIIP 623

Query: 550 HGDHHH 555
              H DH+H
Sbjct: 624 HYDHYH 629
```

 Score = 66.6 bits (161), Expect = 2e-09
 Identities = 46/184 (25%), Positives = 76/184 (41%), Gaps = 29/184 (15%)

```
Query: 144 ADNVRTKDEIN------RQKQEHVKDNEKV---NSNVAVARSQGRYTTNDGYVFNPADII 194
              +D V+  D+I       R +   K N ++   +   VA+ G+YTT DGY+F+P DI
Sbjct: 470 SDKVKLVDDILAFLAPIRHPERLGKPNAQITYTDDEIQVAKLAGKYTTEDGYIFDPRDIT 529

Query: 195 EDTGNAYIVPHGGHYHYIPXXXXXXXXXXXXXXXXXXXXXXXNMQPSQLSYSSTASDNNTQSVA 254
              D G+AY+ PH  H H+I                     + P    ST  ++  + A
Sbjct: 530 SDEGDAYVTPHMTHSHWIKKDSLSEAERAAAQAYAKEKGLTP-----PSTDHQDSGNTEA 584

Query: 255 KGSTSKPANKSENLQSLLKELYDSPSAQRYSESDGLVFDPAKIISRTPNGVAIPHGDHYH 314
              KG+               + +Y+   A +    D + ++       +    IPH DHYH
Sbjct: 585 KGA--------------EAIYNRVKAAKKVPLDRMPYNLQYTVEVKNGSLIIPHYDHYH 629

Query: 315 FIPY 318
              I +
Sbjct: 630 NIKF 633
```

 Score = 53.5 bits (127), Expect = 2e-05
 Identities = 58/302 (19%), Positives = 109/302 (35%), Gaps = 84/302 (27%)

```
Query: 297 IISRTPNGVAIPHGDHYHF----IPYSKLSALEEKIARRVPISGTGSTVSTNAKPNEVVX 352
              +I  T G   HGDHYH+    +PY  + +
Sbjct: 52  VIKITDQGYVTSHGDHYHYYNGKVPYDAIIS--------------------------- 82

Query: 353 XXXXXXXXXXXXXXXXKELSSASDGYIFNPKDIVEETATAYIVR-HGDHFHYIPKS----- 406
                            +EL     Y      DIV E    Y+++  G ++ Y+  +
Sbjct: 83  --------------EELLMKDPNYQLKDSDIVNEIKGGYVIKVDGKYYVYLKDAAHADN 127

Query: 407 --------NQIGQPTLPNNSLATPSPSLPINPGISHEKHEEDGYGFDANRIIAEDESGFI 458
                      Q + + +NS A + +     G        +DGY F+A+ II +    +I
Sbjct: 128 IRTKEEIKRQKQERSHNHNSRADNAVAAARAQG---RYTTDDGYIFNASDIIEDTGDAYI 184

Query: 459 MSHGNHNHYFFKKDLTEEQIKAAQKHLE-------EVKTSHNGLDSLSSHEQDY------ 505
              + HGNH HY  K DL+  ++ AAQ +          +SHN   +      +++
Sbjct: 185 VPHGNHFHYIPKSDLSASELAAAQAYWNGKQGSRPSSSSSHNANPAQPRLSENHNLTVTP 244

Query: 506 PGNAKEMKDLDKKIEEKIAGIMKQYGVKRESI------VVNKEKNAIIYPHGDHHHADPI 559
               +  + +++  + E  A + + V+ + +         + ++     + PHG+H+H  P
Sbjct: 245 TYHQNQGENISSLLRELYAKPLSERHVESDGLIFDPAQITSRTARGVAVPHGNHYHFIPY 304

Query: 560 DE 561
              +
Sbjct: 305 SQ 306
```

 Score = 44.7 bits (104), Expect = 0.009
 Identities = 45/186 (24%), Positives = 63/186 (33%), Gaps = 36/186 (19%)

```
Query:  372  SASDGYIFNPKDIVEETATAYIVRHGDHFHYIPKSN------QIGQPTLPNNSLATPSPS 425
             +  DGYIF+P+DI  +   AY+  H  H H+I K +         Q        L  PS
Sbjct:  516  TTEDGYIFDPRDITSDEGDAYVTPHMTHSHWIKKDSLSEAERAAAQAYAKEKGLTPPSTD 575

Query:  426  LPINPGISHEKHEE--------------DGYGFDANRIIAEDESGFIMSHGNHNH----- 466
             +  G +  K  E                    D  ++    +     I+ H +H H
Sbjct:  576  HQ-DSGNTEAKGAEAIYNRVKAAKKVPLDRMPYNLQYTVEVKNGSLIIPHYDHYHNIKFE 634

Query:  467  ------YFFKKDLTEEQIKAAQK----HLEEVKTSHNGLDSLSSHEQDYPGNAKEMKDLD 516
                   Y   K  TE +A K      H  E   SNG  + S H Q        +    +
Sbjct:  635  WFDEGLYEAPKGYTLEDLLATVKYYVEHPNERPHSDNGFGNASDHVQRNKNGQADTNQTE 694

Query:  517  KKIEEK 522
             K  EEK
Sbjct:  695  KPSEEK 700
```

```
 Score = 39.7 bits (91), Expect = 0.30
 Identities = 34/138 (24%), Positives = 51/138 (36%), Gaps = 19/138 (13%)

Query:  274  ELYDSPSAQRYSESDGLVFDPAKIISRTPNGVAIPHGDHYHFIPYSKLSALEEKIARRV- 332
             E+  +  A +Y+  DG +FDP   I S   +    PH   H H+I   LS  E   A+
Sbjct:  505  EIQVAKLAGKYTTEDGYIFDPRDITSDEGDAYVTPHMTHSHWIKKDSLSEAERAAAQAYA 564

Query:  333  -------PISGTGSTVSTNAKPNEVVXXXXXXXXXXXXXXXXXXXXKELSSASDGYIFNPKDIV 385
                    P +    + +T AK   E +                       D  +N +  V
Sbjct:  565  KEKGLTPPSTDHQDSGNTEAKGAEAIYNRVKAAK----------KVPLDRMPYNLQYTV 613

Query:  386  EETATAYIVRHGDHFHYI 403
             E    + I+ H DH+H I
Sbjct:  614  EVKNGSLIIPHYDHYHNI 631
```

```
tr Q6WNQ0  Surface protein BVH-11 (Fragment) [bvh-11] [Streptococcus  811 AA
           pneumoniae]
                                                                 align
```

```
 Score =  405 bits (1040), Expect = e-111
 Identities = 203/339 (59%), Positives = 253/339 (73%), Gaps = 13/339 (3%)

Query:   21  CAYALNQHRS-QENKDNNRVSYVDGSQSSQKSENLTPDQVSQKEGIQAEQIVIKITDQGY 79
             C+Y L +H++ Q+ K++NRV+Y+DG Q+ QK+ENLTPD+VS++EGI AEQIVIKITDQGY
Sbjct:    1  CSYELGRHQAGQDKKESNRVAYIDGDQAGQKAENLTPDEVSKREGINAEQIVIKITDQGY 60

Query:   80  VTSHGDHYHYYNGKVPYDALFSEELLMKDPNYQLKDADIVNEVKGGYIIKVDGKYYVYLK 139
             VTSHGDHYHYYNGKVPYDA+ SEELLMKDPNYQLKD+DIVNE+KGGY+IKV+GKYYVYLK
Sbjct:   61  VTSHGDHYHYYNGKVPYDAIISEELLMKDPNYQLKDSDIVNEIKGGYVIKVNGKYYVYLK 120

Query:  140  DAAHADNVRTKDEINRQKQEHVKDNEKVNSN---VAVARSQGRYTTNDGYVFNPADIIED 196
             DAAHADN+RTK+EI RQKQEH   N   SN   V  AR+QGRYTT+DGY+FN +DIIED
Sbjct:  121  DAAHADNIRTKEEIKRQKQEH-SHNHGGGSNDQAVVAARAQGRYTTDDGYIFNASDIIED 179

Query:  197  TGNAYIVPHGGHYHYIPXXXXXXXXXXXXXXXXXXXXXXXNMQPSQLSYSSTASDNNTQ---SV 253
             TG+AYIVPHG H+HYIP              Q S+ S SS+ + N Q   S
Sbjct:  180  TGDAYIVPHGNHFHYIP--KSDLSASELAAAQAYWNGKQGSRPSSSSSHNANPAQPRLSE 237
```

```
Query: 254 AKGSTSKPA---NKSENLQSLLKELYDSPSAQRYSESDGLVFDPAKIISRTPNGVAIPHG 310
               T  P    N+ EN+ SLL+ELY  P ++R+ ESDGL+FDPA+I SRT  GVA+PHG
Sbjct: 238 NHNLTVTPTYHQNQGENISSLLRELYAKPLSERHVESDGLIFDPAQITSRTARGVAVPHG 297


Query: 311 DHYHFIPYSKLSALEEKIARRVPISGTGSTVSTNAKPNE 349
           +HYHFIPY ++S LEE+IAR +P+    +     +++P +
Sbjct: 298 NHYHFIPYEQMSELEERIARIIPLRYRSNHWVPDSRPEQ 336



 Score = 73.9 bits (180), Expect = 1e-11
 Identities = 71/272 (26%), Positives = 106/272 (38%), Gaps = 50/272 (18%)


Query: 283 RYSESDGLVFDPAKIISRTPNGVAIPHGDHYHFIPYSKLSALEEKIARRVPISGTGSTVS 342
           RY+  DG +F+ + II  T +   +PHG+H+H+IP S LSA E   A+       GS  S
Sbjct: 162 RYTTDDGYIFNASDIIEDTGDAYIVPHGNHFHYIPKSDLSASELAAAQAYWNGKQGSRPS 221


Query: 343 T----NAKP---------NEVVXXXXXXXXXXXXXXXXXKELSSA--------SDGYIFNP 381
           +     NA P        N  V                 +EL +         SDG IF+P
Sbjct: 222 SSSSHNANPAQPRLSENHNLTVTPTYHQNQGENISSLLRELYAKPLSERHVESDGLIFDP 281


Query: 382 KDIVEETATAYIVRHGDHFHYIP------------------KSNQIGQPTLPNNSLATP 422
             I    TA   V HG+H+H+IP                  +SN     + P       P
Sbjct: 282 AQITSRTARGVAVPHGNHYHFIPYEQMSELEERIARIIPLRYRSNHWVPDSRPEQPSPQP 341


Query: 423 SPS--LPINPGISHEKHEEDGYGFDANRIIAEDESGFIMSHGNHNHYFFKKDLTEEQIKA 480
           SPS    NP +    ++    +A R + +   G++        + Y  KDL+ E
Sbjct: 342 SPSPQPAPNPQPAPSNPIDEKLVKEAVRKVGD---GYVFEENGVSRYIPAKDLSAETAAG 398


Query: 481 AQKHL-EEVKTSHNGLDSLSSHEQDYPGNAKE 511
              L ++    SH     L + + D P + +E
Sbjct: 399 IDSKLAKQESLSH----KLGTKKTDLPSSDRE 426



 Score = 70.1 bits (170), Expect = 2e-10
 Identities = 105/511 (20%), Positives = 192/511 (37%), Gaps = 86/511 (16%)


Query: 71  VIKITDQGYVTSHGDHYHYYNGKVPYDALFSEELLMKDPNYQLKDADIVNEVKGGYIIKV 130
           +I+ T  Y+  HG+H+HY   +P  L + EL    + K       +
Sbjct: 176 IIEDTGDAYIVPHGNHFHY----IPKSDLSASELAAAQAYWNGKQGSRPSSSSSHNANPA 231


Query: 131 DGKYYVYLKDAAHADNVRTKDEINRQKQEHVKDNEKVNSNVAVARSQGRYTTNDGYVFNP 190
              +       +  N+    ++ + + E++  +  +   A   S+ R+  +DG +F+P
Sbjct: 232 QPRL-------SENHNLTVTPTYHQNQGENI--SSLLRELYAKPLSE-RHVESDGLIFDP 281


Query: 191 ADIIEDTGNAYIVPHGGHYHYIPXXXXXXXXXXXXXXXX---XXXXNMQPSQLSYSSTASD 247
           A I   T     VPHG HYH+IP                  + P           +
Sbjct: 282 AQITSRTARGVAVPHGNHYHFIPYEQMSELEERIARIIPLRYRSNHWVPDSRPEQPSPQP 341


Query: 248 NNTQSVAKGSTSKPANKSENLQSLLKELYDSPSAQRYSESDGLVFDPAKIISRTPNGVAI 307
           + + A     P+N  + + L+KE           DG VF+         NGV+
Sbjct: 342 SPSPQPAPNPQPAPSNPID--EKLVKEAVRKVG-------DGYVFEE--------.--NGVS- 383


Query: 308 PHGDHYHFIPYSKLSA-----LEEKIARRVPISGTGSTVSTNAKP------NEVVXXXXX 356
           +IP   LSA      ++  K+A++  +S      T  T+          N+
Sbjct: 384 ------RYIPAKDLSAETAAGIDSKLAKQESLSHKLGTKKTDLPSSDREFYNKAYDLLAR 437


Query: 357 XXXXXXXXXXXXKELSSASDGYIFNPKDI------VEETATAYI--VRHGDHFHYIPKSNQ 408
```

```
                      +    A D  +    KD+      + E   A++   +RH +            +
Sbjct: 438 IHQDLLDNKGRQVDFEALDNLLERLKDVSSDKVKLVEDILAFLAPIRHPE--------R 488

Query: 409 IGQPTLPNNSLATPSPSLPINPGISHEKHEEDGYGFDANRIIAEDESGFIMSHGNHNHYF 468
           +G+P   N+ +      + +   ++ +   EDGY FD  I +++    ++  H   H+H+
Sbjct: 489 LGKP---NSQITYTDDEIQVAK-LAGKYTTEDGYIFDPRDITSDEGDAYVTPHMTHSHWI 544

Query: 469 FKKDLTEEQIKAAQKHLEEVKTSHNGLDSLSSHEQDYPGN--AKEMKDLDKKIE--EKIA 524
            K L+E +  AAQ + +E      GL   S+  QD  GN  AK + +   +++  +K+
Sbjct: 545 KKDSLSEAERAAAQAYAKE-----KGLTPPSTDHQD-SGNTEAKGAEAIYNRVKAAKKVP 598

Query: 525 GIMKQYGVKRESIVVNKEKNAIIYPHGDHHH 555
               Y ++     V  +  ++I PH DH+H
Sbjct: 599 LDRMPYNLQ---YTVEVKNGSLIIPHYDHYH 626


 Score = 67.0 bits (162), Expect = 2e-09
 Identities = 48/190 (25%), Positives = 80/190 (41%), Gaps = 30/190 (15%)

Query: 138 LKDAAHADNVRTKDEIN------RQKQEHVKDNEKV---NSNVAVARSQGRYTTNDGYVF 188
           LKD + +D V+ ++I      R + K N ++ + + VA+ G+YTT DGY+F
Sbjct: 462 LKDVS-SDKVKLVEDILAFLAPIRHPERLGKPNSQITYTDDEIQVAKLAGKYTTEDGYIF 520

Query: 189 NPADIIEDTGNAYIVPHGGHYHYIPXXXXXXXXXXXXXXXXXXXXXNMQPSQLSYSSTASDN 248
           +P DI  D G+AY+ PH  H H+I                     + P     ST   +
Sbjct: 521 DPRDITSDEGDAYVTPHMTHSHWIKKDSLSEAERAAAQAYAKEKGLTP-----PSTDHQD 575

Query: 249 NTQSVAKGSTSKPANKSENLQSLLKELYDSPSAQRYSESDGLVFDPAKIISRTPNGVAIP 308
           +   + AKG+                 + +Y+   A +    D + ++    +      + IP
Sbjct: 576 SGNTEAKGA--------------EAIYNRVKAAKKVPLDRMPYNLQYTVEVKNGSLIIP 620

Query: 309 HGDHYHFIPY 318
           H DHYH I +
Sbjct: 621 HYDHYHNIKF 630


 Score = 53.9 bits (128), Expect = 2e-05
 Identities = 55/300 (18%), Positives = 107/300 (35%), Gaps = 79/300 (26%)

Query: 297 IISRTPNGVAIPHGDHYHF----IPYSKLSALEEKIARRVPISGTGSTVSTNAKPNEVVX 352
           +I  T  G    HGDHYH+    +PY  + +
Sbjct: 52  VIKITDQGYVTSHGDHYHYYNGKVPYDAIIS---------------------------- 82

Query: 353 XXXXXXXXXXXXXXXXKELSSASDGYIFNPKDIVEETATAYIVR-HGDHFHYIPKSNQIGQ 411
                           +EL     Y     DIV E   Y+++ +G ++ Y+  +
Sbjct: 83  --------------EELLMKDPNYQLKDSDIVNEIKGGYVIKVNGKYYVYLKDAAHADN 127

Query: 412 PTLPNNSLATPSPSLPINPGISHEKH-----------EEDGYGFDANRIIAEDESGFIMS 460
           + G S+++             +DGY F+A+ II +      +I+
Sbjct: 128 IRTKEEIKRQKQEHSHNGGGSNDQAVVAARAQGRYTTDDGYIFNASDIIEDTGDAYIVP 187

Query: 461 HGNHNHYFFKKDLTEEQIKAAQKHLE-------EVKTSHNGLDSLSSHEQDY------PG 507
           HGNH HY K DL+ ++ AAQ +        +SHN   +     +++
Sbjct: 188 HGNHFHYIPKSDLSASELAAAQAYWNGKQGSRPSSSSSHNANPAQPRLSENHNLTVTPTY 247

Query: 508 NAKEMKDLDKKIEEKIAGIMKQYGVKRESI------VVNKEKNAIIYPHGDHHHADPIDE 561
           + + +++   + E A + V+ + +      + ++    + PHG+H+H  P ++
Sbjct: 248 HQNQGENISSLLRELYAKPLSERHVESDGLIFDPAQITSRTARGVAVPHGNHYHFIPYEQ 307
```

```
Score = 43.9 bits (102), Expect = 0.016
Identities = 62/286 (21%), Positives = 95/286 (32%), Gaps = 43/286 (15%)

Query: 372 SASDGYIFNPKDIVEETATAYIVRHGDHFHYIPKSN------QIGQPTLPNNSLATPSPS 425
             + DGYIF+P+DI +   AY+ H H H+I K +       Q    L  PS
Sbjct: 513 TTEDGYIFDPRDITSDEGDAYVTPHMTHSHWIKKDSLSEAERAAAQAYAKEKGLTPPSTD 572

Query: 426 LPINPGISHEKHEE--------------DGYGFDANRIIAEDESGFIMSHGNHNH----- 466
             + G + K E              D ++ +     I+ H +H H
Sbjct: 573 HQ-DSGNTEAKGAEAIYNRVKAAKKVPLDRMPYNLQYTVEVKNGSLIIPHYDHYHNIKFE 631

Query: 467 ------YFFKKDLTEEQIKAAQK----HLEEVKTSHNGLDSLSSHEQDYPGNAKEMKDLD 516
                  Y  K T E +A K    H E  S NG  + S H +    +     D D
Sbjct: 632 WFDEGLYEAPKGYTLEDLLATVKYYVEHPNERPHSDNGFGNASDHVRKNKADQDSKPDED 691

Query: 517 KKIEEKIAGIMKQYGVKRESIVVNKEKNAIIYPHGDHHHA-----DPIDEHKPVGIGHSH 571
            K +E   +   K    +N + + P D      D DE +   + HS
Sbjct: 692 KGHDEVSEPTHPESDEKENHAGLNPSADNLYKPSDTEETEEEAEDTTDEAEIPQVEHSV 751

Query: 572 SNYELFKPEEGVAKKEGNKVYTG--EELTNVVNLLKNSTFNNQNFT 615
            N ++  E + K   +    E LT + + L   T +N  +
Sbjct: 752 INAKIADAEALLEKVTDPSIRQNAMETLTGLKSSLLLGTKDNNTIS 797


Score = 39.7 bits (91), Expect = 0.30
Identities = 34/138 (24%), Positives = 51/138 (36%), Gaps = 19/138 (13%)

Query: 274 ELYDSPSAQRYSESDGLVFDPAKIISRTPNGVAIPHGDHYHFIPYSKLSALEEKIARRV- 332
            E+ + A +Y+ DG +FDP  I S   +   PH H H+I  LS E  A+
Sbjct: 502 EIQVAKLAGKYTTEDGYIFDPRDITSDEGDAYVTPHMTHSHWIKKDSLSEAERAAAQAYA 561

Query: 333 -------PISGTGSTVSTNAKPNEVVXXXXXXXXXXXXXXXXXXKELSSASDGYIFNPKDIV 385
                   P +   + +T AK  E +                    D  +N + V
Sbjct: 562 KEKGLTPPSTDHQDSGNTEAKGAEAIYNRVKAAK----------KVPLDRMPYNLQYTV 610

Query: 386 EETATAYIVRHGDHFHYI 403
            E    + I+ H DH+H I
Sbjct: 611 EVKNGSLIIPHYDHYHNI 628
```

tr <u>Q6WNP6</u> **Surface protein BVH-11-2 (Fragment) [bvh-11-2] [Streptococcus** 816 AA
            pneumoniae]

                                                                        <u>align</u>

```
Score =  400 bits (1028), Expect = e-110
Identities = 202/339 (59%), Positives = 251/339 (73%), Gaps = 13/339 (3%)

Query: 21 CAYALNQHRS-QENKDNNRVSYVDGSQSSQKSENLTPDQVSQKEGIQAEQIVIKITDQGY 79
           C+Y L +H++ Q+ K++NRV+Y+DG Q+ QK+ENLTPD+VS++EGI AEQIVIKITDQGY
Sbjct: 1  CSYELGRHQAGQDKKESNRVAYIDGDQAGQKAENLTPDEVSKREGINAEQIVIKITDQGY 60

Query: 80 VTSHGDHYHYYNGKVPYDALFSEELLMKDPNYQLKDADIVNEVKGGYIIKVDGKYYVYLK 139
           VTSHGDHYHYYNGKVPYDA+ SEELLMKDPNYQLKD+DIVNE+KGGY+IKV+GKYYVYLK
Sbjct: 61 VTSHGDHYHYYNGKVPYDAIISEELLMKDPNYQLKDSDIVNEIKGGYVIKVNGKYYVYLK 120
```

```
Query:  140  DAAHADNVRTKDEINRQKQEHVKDNEKVNSN---VAVARSQGRYTTNDGYVFNPADIIED  196
             DAAHADN+RTK+EI RQKQEH   N   SN   V AR+QGRYTT+DGY+FN +DIIED
Sbjct:  121  DAAHADNIRTKEEIKRQKQEH-SHNHGGGSNDQAVVAARAQGRYTTDDGYIFNASDIIED  179

Query:  197  TGNAYIVPHGGHYHYIPXXXXXXXXXXXXXXXXXXXNMQPSQLSYSSTASDNNTQ---SV  253
             TG+AYIVP G H+HYIP                   Q S+ S SS+ + N  Q   S
Sbjct:  180  TGDAYIVPRGNHFHYIP--KSDLSASELAAAQAYWNGKQGSRPSSSSSHNANPAQPRLSE  237

Query:  254  AKGSTSKPA---NKSENLQSLLKELYDSPSAQRYSESDGLVFDPAKIISRTPNGVAIPHG  310
                 T P    N+ EN+ SLL+ELY  P ++R  ESDGL+FDPA+I SRT  GVA+PHG
Sbjct:  238  NHNLTVTPTYHQNQGENISSLLRELYAKPLSERRVESDGLIFDPAQITSRTARGVAVPHG  297

Query:  311  DHYHFIPYSKLSALEEKIARRVPISGTGSTVSTNAKPNE  349
             +HYHFIPY ++S LEE+IAR +P+    +   +++P +
Sbjct:  298  NHYHFIPYEQMSELEERIARIIPLRYRSNHWVPDSRPEQ  336


  Score = 70.9 bits (172), Expect = 1e-10
  Identities = 70/272 (25%), Positives = 105/272 (37%), Gaps = 50/272 (18%)

Query:  283  RYSESDGLVFDPAKIISRTPNGVAIPHGDHYHFIPYSKLSALEEKIARRVPISGTGSTVS  342
             RY+  DG +F+ + II  T +    +P G+H+H+IP S LSA E   A+      GS  S
Sbjct:  162  RYTTDDGYIFNASDIIEDTGDAYIVPRGNHFHYIPKSDLSASELAAAQAYWNGKQGSRPS  221

Query:  343  T----NAKP---------NEVVXXXXXXXXXXXXXXXXXKELSSA--------SDGYIFNP  381
             +    NA P         N V               +EL +        SDG IF+P
Sbjct:  222  SSSSHNANPAQPRLSENHNLTVTPTYHQNQGENISSLLRELYAKPLSERRVESDGLIFDP  281

Query:  382  KDIVEETATAYIVRHGDHFHYIP------------------KSNQIGQPTLPNNSLATP  422
              I   TA   V HG++H+IP                  +SN   + P     P
Sbjct:  282  AQITSRTARGVAVPHGNHYHFIPYEQMSELEERIARIIPLRYRSNHWVPDSRPEQPSPQP  341

Query:  423  SPS--LPINPGISHEKHEEDGYGFDANRIIAEDESGFIMSHGNHNHYFFKKDLTEEQIKA  480
             SPS    NP +   ++    +A R + +   G++       + Y   KDL+ E
Sbjct:  342  SPSPQPAPNPQPAPSNPIDEKLVKEAVRKVGD---GYVFEENGVSRYIPAKDLSAETAAG  398

Query:  481  AQKHL-EEVKTSHNGLDSLSSHEQDYPGNAKE  511
                 L ++   SH    L + + D P + +E
Sbjct:  399  IDSKLAKQESLSH----KLGTKKTDLPSSDRE  426


  Score = 67.0 bits (162), Expect = 2e-09
  Identities = 48/190 (25%), Positives = 80/190 (41%), Gaps = 30/190 (15%)

Query:  138  LKDAAHADNVRTKDEIN------RQKQEHVKDNEKV---NSNVAVARSQGRYTTNDGYVF  188
             LKD + +D V+ +++I      R +  K N ++    +   + VA+  G+YTT DGY+F
Sbjct:  462  LKDVS-SDKVKLVEDILAFLAPIRHPERLGKPNSQITYTDDEIQVAKLAGKYTTEDGYIF  520

Query:  189  NPADIIEDTGNAYIVPHGGHYHYIPXXXXXXXXXXXXXXXXXXXXXXXNMQPSQLSYSSTASDN  248
             +P DI  D G+AY+ PH  H H+I                    +   P      ST   +
Sbjct:  521  DPRDITSDEGDAYVTPHMTHSHWIKKDSLSEAERAAAQAYAKEKGLTP-----PSTDHQD  575

Query:  249  NTQSVAKGSTSKPANKSENLQSLLKELYDSPSAQRYSESDGLVFDPAKIISRTPNGVAIP  308
             + + AKG+            + +Y+    A +   D + ++       +       + IP
Sbjct:  576  SGNTEAKGA--------------EAIYNRVKAAKKVPLDRMPYNLQYTVEVKNGSLIIP  620

Query:  309  HGDHYHFIPY  318
```

```
                H DHYH I +
Sbjct: 621 HYDHYHNIKF 630


   Score = 50.8 bits (120), Expect = 1e-04
   Identities = 54/300 (18%), Positives = 106/300 (35%), Gaps = 79/300 (26%)

Query: 297 IISRTPNGVAIPHGDHYHF----IPYSKLSALEEKIARRVPISGTGSTVSTNAKPNEVVX 352
            +I  T G     HGDHYH+    +PY + +
Sbjct: 52  VIKITDQGYVTSHGDHYHYYNGKVPYDAIIS--------------------------- 82

Query: 353 XXXXXXXXXXXXXXXXXKELSSASDGYIFNPKDIVEETATAYIVR-HGDHFHYIPKSNQIGQ 411
                             +EL     Y     DIV E     Y+++ +G ++ Y+  +
Sbjct: 83  ---------------EELLMKDPNYQLKDSDIVNEIKGGYVIKVNGKYYVYLKDAAHADN 127

Query: 412 PTLPNNSLATPSPSLPINPGISHEKH-----------EEDGYGFDANRIIAEDESGFIMS 460
                        + G S+++              +DGY F+A+ II +    +I+
Sbjct: 128 IRTKEEIKRQKQEHSNHGGGSNDQAVVAARAQGRYTTDDGYIFNASDIIEDTGDAYIVP 187

Query: 461 HGNHNHYFFKKDLTEEQIKAAQKHLE-------EVKTSHNGLDSLSSHEQDY------PG 507
            GNH HY  K DL+  ++ AAQ +       +SHN    +      +++
Sbjct: 188 RGNHFHYIPKSDLSASELAAAQAYWNGKQGSRPSSSSSHNANPAQPRLSENHNLTVTPTY 247

Query: 508 NAKEMKDLDKKIEEKIAGIMKQYGVKRESI------VVNKEKNAIIYPHGDHHHADPIDE 561
            +  + +++    + E  A  + +  V+ + +         + ++    + PHG+H+H  P ++
Sbjct: 248 HQNQGENISSLLRELYAKPLSERRVESDGLIFDPAQITSRTARGVAVPHGNHYHFIPYEQ 307


   Score = 43.9 bits (102), Expect = 0.016
   Identities = 62/286 (21%), Positives = 95/286 (32%), Gaps = 43/286 (15%)

Query: 372 SASDGYIFNPKDIVEETATAYIVRHGDHFHYIPKSN------QIGQPTLPNNSLATPSPS 425
            +  DGYIF+P+DI +   AY+ H H H+I K +     Q      L  PS
Sbjct: 513 TTEDGYIFDPRDITSDEGDAYVTPHMTHSHWIKKDSLSEAERAAAQAYAKEKGLTPPSTD 572

Query: 426 LPINPGISHEKHEE-------------DGYGFDANRIIAEDESGFIMSHGNHNH----- 466
            + G + K E             D  ++    + I+ H +H H
Sbjct: 573 HQ-DSGNTEAKGAEAIYNRVKAAKKVPLDRMPYNLQYTVEVKNGSLIIPHYDHYHNIKFE 631

Query: 467 ------YFFKKDLTEEQIKAAQK----HLEEVKTSHNGLDSLSSHEQDYPGNAKEMKDLD 516
                  Y  K TE+ A K    H E   S NG  + S H +   +    D D
Sbjct: 632 WFDEGLYEAPKGYTLEDLLATVKYYVEHPNERPHSDNGFGNASDHVRKNKADQDSKPDED 691

Query: 517 KKIEEKIAGIMKQYGVKRESIVVNKEKNAIIYPHGDHHHA-----DPIDEHKPVGIGHSH 571
            K +E     + K   +N  + +  P D    D DE +  + HS
Sbjct: 692 KGHDEVSEPTHPESDEKENHAGLNPSADNLYKPSTDTEETEEEAEDTTDEAEIPQVEHSV 751

Query: 572 SNYELFKPEEGVAKKEGNKVYTG--EELTNVVNLLKNSTFNNQNFT 615
            N ++   E  + K   +      E LT + + L   T +N  +
Sbjct: 752 INAKIADAEALLEKVTDPSIRQNAMETLTGLKSSLLLGTKDNNTIS 797


   Score = 39.7 bits (91), Expect = 0.30
   Identities = 34/138 (24%), Positives = 51/138 (36%), Gaps = 19/138 (13%)

Query: 274 ELYDSPSAQRYSESDGLVFDPAKIISRTPNGVAIPHGDHYHFIPYSKLSALEEKIARRV- 332
            E+ + A +Y+  DG +FDP I S  +   PH  H H+I   LS  E    A+
Sbjct: 502 EIQVAKLAGKYTTEDGYIFDPRDITSDEGDAYVTPHMTHSHWIKKDSLSEAERAAAQAYA 561
```

```
Query: 333 -------PISGTGSTVSTNAKPNEVVXXXXXXXXXXXXXXXXXXKELSSASDGYIFNPKDIV 385
                  P +    + +T AK  E +                        D  +N +  V
Sbjct: 562 KEKGLTPPSTDHQDSGNTEAKGAEAIYNRVKAAK----------KVPLDRMPYNLQYTV 610


Query: 386 EETATAYIVRHGDHFHYI 403
               E    + I+ H DH+H I
Sbjct: 611 EVKNGSLIIPHYDHYHNI 628
```


tr <u>Q8NZ82</u> **Hypothetical protein spyM18_2072 [spyM18_2072] [Streptococcus**   823 AA
               pyogenes (serotype M18)]

                                                                                    <u>align</u>


```
 Score =  306 bits (785), Expect = 1e-81
 Identities = 198/612 (32%), Positives = 299/612 (48%), Gaps = 74/612 (12%)

Query: 1   MKFSKKYIAAGSAVIVSLSLCAYALNQHRSQENKDNNRVSYVDGSQSSQKSE--NLTPDQ 58
           MK +   YI + +A++++  + +Y L +H         +N+++Y+D S+    K+   N T DQ
Sbjct: 1   MKKTYGYIGSVAAILLATHIGSYQLGKHHMGPVTKDNQIAYIDDSKGKAKAPKTNKTMDQ 60

Query: 59  VSQKEGIQAEQIVIKITDQGYVTSHGDHYHYYNGKVPYDALFSEELLMKDPNYQLKDADI 118
           +S +EGI AEQIV+KITDQGYVTSHGDHYH+YNGKVPYDA+ SEELLM DPNY+ K +D+
Sbjct: 61  ISAEEGISAEQIVVKITDQGYVTSHGDHYHFYNGKVPYDAIISEELLMTDPNYRFKQSDV 120

Query: 119 VNEVKGGYIIKVDGKYYVYLKDAAHADNVRTKDEINRQKQEHVKD-NEKVNSNVA----- 172
           +NE+  GY+IKV+G YYVYLK +    N+RTK +I  Q  +  K+  EK   + VA
Sbjct: 121 INEILDGYVIKVNGNYYVYLKPGSKRKNIRTKQQIAEQVAKGTKEAKEKGLAQVAHLSKE 180

Query: 173 ------VARSQGRYTTNDGYVFNPADIIEDTGNAYIVPHGGHYHYIPXXXXXXXXXXXXXX 226
                 A+ QGRYTT+DGY+F+P DII+D G+AY+VPHG HYHYIP
Sbjct: 181 EVAAVNEAKRQGRYTTDDGYIFSPTDIIDDLGDAYLVPHGNHYHYIPKKDLSPSELAAAQ 240

Query: 227 XXXXXXNMQPSQLS-YSSTASDNNTQSVAKGSTSKPAN-------------------- 263
                 + ++ S Y  T +    ++     T  P
Sbjct: 241 AYWSQKQGRGARPSDYRPTPAPGRRKAPIPDVTPNPRQGHQPDNGGYHPAPPRPNDASQN 300

Query: 264 -------KSENLQSLLKELYDSPSAQRYSESDGLVFDPAKIISRTPNGVAIPHGDHYHFI 316
                  K +   + LL +L+       R+ E DGL+F+P ++I   G   +PHGDHYH I
Sbjct: 301 KHQRDEFKGKTFKELLDQLHRLDLKYRHVEEDGLIFEPTQVIKSNAFGYVVPHGDHYHII 360

Query: 317 PYSKLSALEEKIARRVPISGT--GSTVSTNAKPN--EVVXXXXXXXXXXXXXXXXXXKELSS 372
           P S+LS LE ++A R      T    + S ++KP+  EV                   +
Sbjct: 361 PRSQLSPLEMELADRYLAGQTEDNDSGSDHSKPSDKEVTHTFLGHRIKAYGKGLDGKPYD 420

Query: 373 ASDGYIFNPKDIVEETATAYIVRHGDHFHYIPKSNQIGQPTLPN--NSLATPSPSLPINP 430
           SD Y+F+ + I     +   +HGDHFHYI    ++ Q  L   N +      +   +
Sbjct: 421 TSDAYVFSKESIHSVDKSGVTAKHGDHFHYI-GFGELEQYELDEVANWVKAKGQADELAA 479

Query: 431 GISHEKHEEDGYGFD----ANRIIAEDESGFIMSHGNHNHYYFFKKDLTEEQIKAAQKHLE 486
           +  E+ +E     FD    + ++  + + G+IM    ++++ +  L    QI  A++ L
Sbjct: 480 ALDQEQGKEKPL-FDTKKVSRKVTKDGKVGYIMPKDGKDYFYARDQLDLTQIAFAEQELM 538

Query: 487 EVKTSHNGLDSLSSHEQDYPGNAKEMKDLDKKIEEKIAGIMKQYGVKRESIVVNKEKNAI 546
               H   D                 +D  IE ++A +    +    +   +  + + +
Sbjct: 539 LKDKKHYRYDI----------------VDTGIEPRLAVDVSSLPMHAGNATYDTGSSFV 581
```

```
Query:  547  IYPHGDHHHADP 558
             I PH DH H  P
Sbjct:  582  I-PHIDHIHVVP 592
```

 Score = 51.2 bits (121), Expect = 1e-04
 Identities = 54/313 (17%), Positives = 109/313 (34%), Gaps = 35/313 (11%)

```
Query:   67  AEQIVIKITDQGYVTSHGDHYHY--YNGKVPYD-----------------ALFSEELLM 106
             +++ +  +   G   HGDH+HY +    Y+                    A   +E
Sbjct:  428  SKESIHSVDKSGVTAKHGDHFHYIGFGELEQYELDEVANWVKAKGQADELAAALDQEQGK 487

Query:  107  KDPNYQLKDAD--IVNEVKGGYIIKVDGKYYVYLKDAAHADNVRTKDEINRQKQEHVKDN 164
             + P + K   +  + K GYI+  DGK Y Y +D     +   ++    +   +KD
Sbjct:  488  EKPLFDTKKVSRKVTKDGKVGYIMPKDGKDYFYARDQLDLTQIAFAEQ-----ELMLKDK 542

Query:  165  EKVNSNVAVARSQGRYTTN-DGYVFNPADIIEDTGNAYIVPHGGHYHYIPXXXXXXXXXX 223
             +      ++     + R +     + +      DTG++++++PH  H H +P
Sbjct:  543  KHYRYDIVDTGIEPRLAVDVSSLPMHAGNATYDTGSSFVIPHIDHIHVVPYSWLTRDQIA 602

Query:  224  XXXXXXXXXNMQPSQLSYSSTASDNNTQSVAKGSTSKPANKSENLQSLLKELYDSPSAQR 283
                      ++P   S    +         +    +   +   +    +E+  + +  R
Sbjct:  603  TIKYVMQHPEVRPDVWSKPGHEESGSVIPNVTPLDKRAGMPNWQIIHSAEEVQKALAEGR 662

Query:  284  YSESDGLVFDPAKIISR-----TPNGVAIPH--GDHYHFIPYSKLSALEEKIARRVPISG 336
             ++   DG +FDP ++++     +IP   G       I  S LS  E + A+ +
Sbjct:  663  FATPDGYIFDPRDVLAKETFVWKDGSFSIPRADGSSLRTINKSDLSQAEWQQAQELLAKK 722

Query:  337  TGSTVSTNAKPNE 349
                 +  KP E
Sbjct:  723  NAGDATDTDKPKE 735
```

 Score = 38.1 bits (87), Expect = 0.88
 Identities = 39/199 (19%), Positives = 78/199 (38%), Gaps = 33/199 (16%)

```
Query:  372  SASDGYIFNPKDIVEETATAYIVRHGDHFHYIPKSNQIGQPTLPNNSLATPSPSLPINPG 431
             SA +G I + +V+ T   Y+  HGDH+H+                    +  +P +
Sbjct:   62  SAEEG-ISAEQIVVKITDQGYVTSHGDHYHFY-------------------NGKVPYDAI 101

Query:  432  ISHEKHEED-GYGFDANRIIAEDESGFIMSHGNHNHYFFKKDLTEEQIKAAQKHLEEV-- 488
             IS E    D Y F + +I E  G+++    + + K      + I+  Q+  E+V
Sbjct:  102  ISEELLMTDPNYRFKQSDVINEILDGYVIKVNGNYYVYLKPGSKRKNIRTKQQIAEQVAK 161

Query:  489  ---KTSHNGLDSLSSHEQDYPGNAKEMKDLDKKIEEKIAGIMKQYGVKRESIVVNKEKNA 545
                +   GL  ++   ++       E K  +  +         +    +  +++   +A
Sbjct:  162  GTKEAKEKGLAQVAHLSKEEVAAVNEAKRQGRYTTD-------DGYIFSPTDIIDDLGDA 214

Query:  546  IIYPHGDHHHADPIDEHKP 564
             + PHG+H+H  P + P
Sbjct:  215  YLVPHGNHYHYIPKKDLSP 233
```


 tr <u>Q8E4U1</u> **Hypothetical protein gbs1306 [gbs1306] [Streptococcus   822
                 agalactiae**   AA

(serotype III)]

```
Score =  305 bits (782), Expect = 2e-81
Identities = 197/612 (32%), Positives = 299/612 (48%), Gaps = 74/612 (12%)

Query: 1    MKFSKKYIAAGSAVIVSLSLCAYALNQHRSQENKDNNRVSYVDGSQSSQKSE--NLTPDQ 58
            MK +  YI + +A++++   + +Y L +H         +N+++Y+D S+   K+    N T DQ
Sbjct: 1    MKKTYGYIGSVAAILLATHIGSYQLGKHHMGLATKDNQIAYIDDSKGKAKAPKTNKTMDQ 60

Query: 59   VSQKEGIQAEQIVIKITDQGYVTSHGDHYHYYNGKVPYDALFSEELLMKDPNYQLKDADI 118
            +S +EGI AEQIV+KITDQGYVTSHGDHYH+YNGKVPYDA+ SEELLM DPNY+ K +D+
Sbjct: 61   ISAEEGISAEQIVVKITDQGYVTSHGDHYHFYNGKVPYDAIISEELLMTDPNYRFKQSDV 120

Query: 119  VNEVKGGYIIKVDGKYYVYLKDAAHADNVRTKDEINRQKQEHVKD-NEKVNSNVA----- 172
            +NE+  GY+IKV+G YYVYLK  +    N+RTK +I  Q  + K+  EK + VA
Sbjct: 121  INEILDGYVIKVNGNYYVYLKPGSKRKNIRTKQQIAEQVAKGTKEAKEKGLAQVAHLSKE 180

Query: 173  ------VARSQGRYTTNDGYVFNPADIIEDTGNAYIVPHGGHYHYIPXXXXXXXXXXXXX 226
                  A+ QGRYTT+DGY+F+P DII+D G+AY+VPHG HYHYIP
Sbjct: 181  EVAAVNEAKRQGRYTTDDGYIFSPTDIIDDLGDAYLVPHGNHYHYIPKKDLSPSELAAAQ 240

Query: 227  XXXXXXNMQPSQLS-YSSTASDNNTQSVAKGSTSKPAN-------------------- 263
                  + ++ S Y  T +    ++    T  P
Sbjct: 241  AYWSQKQGRGARPSDYRPTPAPGRRKAPIPDVTPNPGQGHQPDNGGYHPAPPRPNDASQN 300

Query: 264  -------KSENLQSLLKELYDSPSAQRYSESDGLVFDPAKIISRTPNGVAIPHGDHYHFI 316
                   K +  + LL +L+        R+ E DGL+F+P ++I    G +PHGDHYH I
Sbjct: 301  KHQRDEFKGKTFKELLDQLHRLDLKYRHVEEDGLIFEPTQVIKSNAFGYVVPHGDHYHII 360

Query: 317  PYSKLSALEEKIARRVPISGT--GSTVSTNAKPN--EVVXXXXXXXXXXXXXXXXKELSS 372
            P S+LS LE ++A R     T      + S ++KP+  EV                   +
Sbjct: 361  PRSQLSPLEMELADRYLAGQTEDNDSGSEHSKPSDKEVTHTFLGHRIKAYGKGLDGKPYD 420

Query: 373  ASDGYIFNPKDIVEETATAYIVRHGDHFHYIPKSNQIGQPTLPN--NSLATPSPSLPINP 430
            SD Y+F+ + I     +  +HGDHFHYI    ++ Q L  N +     + +
Sbjct: 421  TSDAYVFSKESIHSVDKSGVTAKHGDHFHYI-GFGELEQYELDEVANWVKAKGQADELAA 479

Query: 431  GISHEKHEEDGYGFD----ANRIIAEDESGFIMSHGNHNHYFFKKDLTEEQIKAAQKHLE 486
            + E+ +E    FD   + ++ + + G++M     ++++ +  L  QI  A++ L
Sbjct: 480  ALDQEQGKEKPL-FDTKKVSRKVTKDGKVGYMMPKDGKDYFYARDQLDLTQIAFAEQELM 538

Query: 487  EVKTSHNGLDSLSSHEQDYPGNAKEMKDLDKKIEEKIAGIMKQYGVKRESIVVNKEKNAI 546
            H   D            +D IE ++A  +   +     +     + +
Sbjct: 539  LKDKKHYRYDI----------------VDTGIEPRLAVDVSSLPMHAGNATYDTGSSFV 581

Query: 547  IYPHGDHHHADP 558
            I PH DH H  P
Sbjct: 582  I-PHIDHIHVVP 592


Score = 49.3 bits (116), Expect = 4e-04
Identities = 53/313 (16%), Positives = 109/313 (33%), Gaps = 35/313 (11%)

Query: 67   AEQIVIKITDQGYVTSHGDHYHY--YNGKVPYD-----------------ALFSEELLM 106
            +++ +  +   G   HGDH+HY  +     Y+                    A  +E
Sbjct: 428  SKESIHSVDKSGVTAKHGDHFHYIGFGELEQYELDEVANWVKAKGQADELAAALDQEQGK 487

Query: 107  KDPNYQLKDAD--IVNEVKGGYIIKVDGKYYVYLKDAAHADNVRTKDEINRQKQEHVKDN 164
```

```
               + P + K      +  + K GY++  DGK Y Y +D     +   ++    +  +KD
Sbjct: 488 EKPLFDTKKVSRKVTKDGKVGYMMPKDGKDYFYARDQLDLTQIAFAEQ-----ELMLKDK 542

Query: 165 EKVNSNVAVARSQGRYTTN-DGYVFNPADIIEDTGNAYIVPHGGHYHYIPXXXXXXXXXX 223
               +    ++     + R  +     +  +  DTG+++++PH H H +P
Sbjct: 543 KHYRYDIVDTGIEPRLAVDVSSLPMHAGNATYDTGSSFVIPHIDHIHVVPYSWLTRDQIA 602

Query: 224 XXXXXXXXXNMQPSQLSYSSTASDNNTQSVAKGSTSKPANKSENLQSLLKELYDSPSAQR 283
                      ++P  S      +         +     +   +  +  +E+  +  R
Sbjct: 603 TVKYVMQHPEVRPDVWSKPGHEESGSVIPNVTPLDKRAGMPNWQIIHSAEEVQKALAEGR 662

Query: 284 YSESDGLVFDPAKIISR-----TPNGVAIPH--GDHYHFIPYSKLSALEEKIARRVPISG 336
               ++  DG +FDP ++++         +IP  G     I  S LS  E + A+ +
Sbjct: 663 FATPDGYIFDPRDVLAKETFVWKDGSFSIPRADGSSLRTINKSDLSQAEWQQAQELLAKK 722

Query: 337 TGSTVSTNAKPNE 349
                +   KP E
Sbjct: 723 NTGDATDTDKPKE 735


  Score = 38.1 bits (87), Expect = 0.88
  Identities = 39/199 (19%), Positives = 78/199 (38%), Gaps = 33/199 (16%)

Query: 372 SASDGYIFNPKDIVEETATAYIVRHGDHFHYIPKSNQIGQPTLPNNSLATPSPSLPINPG 431
           SA +G I  + +V+ T   Y+  HGDH+H+                   +   +P +
Sbjct: 62  SAEEG-ISAEQIVVKITDQGYVTSHGDHYHFY------------------NGKVPYDAI 101

Query: 432 ISHEKHEED-GYGFDANRIIAEDESGFIMSHGNHNHYFFKKDLTEEQIKAAQKHLEEV-- 488
           IS E    D  Y F + +I E  G+++   + + + K     + I+ Q+  E+V
Sbjct: 102 ISEELLMTDPNYRFKQSDVINEILDGYVIKVNGNYYVYLKPGSKRKNIRTKQQIAEQVAK 161

Query: 489 ---KTSHNGLDSLSSHEQDYPGNAKEMKDLDKKIEEKIAGIMKQYGVKRESIVVNKEKNA 545
              +   GL  ++   ++        E K   +  +          +    + +++   +A
Sbjct: 162 GTKEAKEKGLAQVAHLSKEEVAAVNEAKRQGRYTTD-------DGYIFSPTDIIDDLGDA 214

Query: 546 IIYPHGDHHHADPIDEHKP 564
           + PHG+H+H  P  + P
Sbjct: 215 YLVPHGNHYHYIPKKDLSP 233



tr   Q8DZ81    Streptococcal histidine triad family protein      822 AA
               [SAG1233]                                          align
               [Streptococcus agalactiae (serotype V)]


  Score = 305 bits (782), Expect = 2e-81
  Identities = 198/612 (32%), Positives = 298/612 (48%), Gaps = 74/612 (12%)

Query: 1   MKFSKKYIAAGSAVIVSLSLCAYALNQHRSQENKDNNRVSYVDGSQSSQKSE--NLTPDQ 58
           MK + YI + +A++++  + +Y L +H      +N+++Y+D S+  K+   N T DQ
Sbjct: 1   MKKTYGYIGSVAAILLATHIGSYQLGKHHMGLATKDNQIAYIDDSKGKVKAPKTNKTMDQ 60

Query: 59  VSQKEGIQAEQIVIKITDQGYVTSHGDHYHYYNGKVPYDALFSEELLMKDPNYQLKDADI 118
           +S +EGI AEQIV+KITDQGYVTSHGDHYH+YNGKVPYDA+ SEELLM DPNY  K +D+
Sbjct: 61  ISAEEGISAEQIVVKITDQGYVTSHGDHYHFYNGKVPYDAIISEELLMTDPNYHFKQSDV 120

Query: 119 VNEVKGGYIIKVDGKYYVYLKDAAHADNVRTKDEINRQKQEHVKD-NEKVNSNVA----- 172
```

```
              +NE+   GY+IKV+G YYVYLK +    N+RTK +I  Q  +  K+  EK  + VA
Sbjct: 121 INEILDGYVIKVNGNYYVYLKPGSKRKNIRTKQQIAEQVAKGTKEAKEKGLAQVAHLSKE 180

Query: 173 ------VARSQGRYTTNDGYVFNPADIIEDTGNAYIVPHGGHYHYIPXXXXXXXXXXXXXX 226
                 A+  QGRYTT+DGY+F+P DII+D G+AY+VPHG HYHYIP
Sbjct: 181 EVAAVNEAKRQGRYTTDDGYIFSPTDIIDDLGDAYLVPHGNHYHYIPKKDLSPSELAAAQ 240

Query: 227 XXXXXXNMQPSQLS-YSSTASDNNTQSVAKGSTSKPAN-------------------- 263
                 + ++ S Y T +    ++       T  P
Sbjct: 241 AYWSQKQGRGARPSDYRPTPAPGRRKAPIPDVTPNPGQGHQPDNGGYHPAPPRPNDASQN 300

Query: 264 -------KSENLQSLLKELYDSPSAQRYSESDGLVFDPAKIISRTPNGVAIPHGDHYHFI 316
                  K +  + LL +L++    R+ E DGL+F+P ++I    G +PHGDHYH I
Sbjct: 301 KHQRDEFKGKTFKELLDQLHRLDLKYRHVEEDGLIFEPTQVIKSNAFGYVVPHGDHYHII 360

Query: 317 PYSKLSALEEKIARRVPISGT--GSTVSTNAKPN--EVVXXXXXXXXXXXXXXXXXKELSS 372
              P S+LS LE ++A R     T   + S ++KP+   EV                 +
Sbjct: 361 PRSQLSPLEMELADRYLAGQTDDNDSGSDHSKPSDKEVTHTFLGHRIKAYGKGLDGKPYD 420

Query: 373 ASDGYIFNPKDIVEETATAYIVRHGDHFHYIPKSNQIGQPTLPN--NSLATPSPSLPINP 430
              SD Y+F+ + I  +   +HGDHFHYI    ++ Q L   N+    + +
Sbjct: 421 TSDAYVFSKESIHSVDKSGVTAKHGDHFHYI-GFGELEQYELDEVANWVKAKGQADELVA 479

Query: 431 GISHEKHEEDGYGFD----ANRIIAEDESGFIMSHGNHNHYFFKKDLTEEQIKAAQKHLE 486
               +  E+ +E    FD    + ++  + + G+IM    ++++ +  L   QI  A++ L
Sbjct: 480 ALDQEQGKEKPL-FDTKKVSRKVTKDGKVGYIMPKDGKDYFYARYQLDLTQIAFAEQELM 538

Query: 487 EVKTSHNGLDSLSSHEQDYPGNAKEMKDLDKKIEEKIAGIMKQYGVKRESIVVNKEKNAI 546
                H    D              +D  IE ++A  +   +     +    +    +  +
Sbjct: 539 LKDKKHYRYDI----------------VDTGIEPRLAVDLSSLPMHAGNATYDTGSSFV 581

Query: 547 IYPHGDHHHADP 558
              I PH DH H  P
Sbjct: 582 I-PHIDHIHVVP 592


   Score = 47.8 bits (112), Expect = 0.001
   Identities = 53/313 (16%), Positives = 108/313 (33%), Gaps = 35/313 (11%)

Query: 67  AEQIVIKITDQGYVTSHGDHYHY--YNGKVPYD-----------------ALFSEELLM 106
              +++ +  +   G    HGDH+HY  +     Y+                  A   +E
Sbjct: 428 SKESIHSVDKSGVTAKHGDHFHYIGFGELEQYELDEVANWVKAKGQADELVAALDQEQGK 487

Query: 107 KDPNYQLKDAD--IVNEVKGGYIIKVDGKYYVYLKDAAHADNVRTKDEINRQKQEHVKDN 164
              + P + K     + + K GYI+ DGK Y Y +      +     + ++    +  +KD
Sbjct: 488 EKPLFDTKKVSRKVTKDGKVGYIMPKDGKDYFYARYQLDLTQIAFAEQ-----ELMLKDK 542

Query: 165 EKVNSNVAVARSQGRYTTN-DGYVFNPADIIEDTGNAYIVPHGGHYHYIPXXXXXXXXXX 223
                +   ++     + R  +    + +       + + +   DTG++++PH  H H +P
Sbjct: 543 KHYRYDIVDTGIEPRLAVDLSSLPMHAGNATYDTGSSFVIPHIDHIHVVPYSWLTRNQIA 602

Query: 224 XXXXXXXXXNMQPSQLSYSSTASDNNTQSVAKGSTSKPANKSENLQSLLKELYDSPSAQR 283
                       ++P   S         +       +   + +   +E+  + + R
Sbjct: 603 TIKYVMQHPEVRPDVWSKPGHEESGSVIPNVTPLDKRAGMPNWQIIHSAEEVQKALAEGR 662

Query: 284 YSESDGLVFDPAKIISR-----TPNGVAIPH--GDHYHFIPYSKLSALEEKIARRVPISG 336
              ++  DG +FDP ++++       +IP  G    I  S LS  E + A+ +
Sbjct: 663 FAAPDGYIFDPRDVLAKETFVWKDGSFSIPRADGSSLRTINKSDLSQAEWQQAQELLAKK 722
```

```
Query: 337 TGSTVSTNAKPNE 349
               +   KP E
Sbjct: 723 NAGDATDTDKPEE 735
```

```
 Score = 38.1 bits (87), Expect = 0.88
 Identities = 39/199 (19%), Positives = 78/199 (38%), Gaps = 33/199 (16%)
```

```
Query: 372 SASDGYIFNPKDIVEETATAYIVRHGDHFHYIPKSNQIGQPTLPNNSLATPSPSLPINPG 431
           SA +G I  + +V+ T   Y+  HGDH+H+                      +   +P +
Sbjct: 62  SAEEG-ISAEQIVVKITDQGYVTSHGDHYHFY-------------------NGKVPYDAI 101
```

```
Query: 432 ISHEKHEED-GYGFDANRIIAEDESGFIMSHGNHNHYFFKKDLTEEQIKAAQKHLEEV-- 488
           IS E    D Y F + +I E  G+++   + + + K    + I+ Q+  E+V
Sbjct: 102 ISEELLMTDPNYHFKQSDVINEILDGYVIKVNGNYYVYLKPGSKRKNIRTKQQIAEQVAK 161
```

```
Query: 489 ---KTSHNGLDSLSSHEQDYPGNAKEMKDLDKKIEEKIAGIMKQYGVKRESIVVNKEKNA 545
              +    GL  ++   ++       E K  +   +            +   + +++   +A
Sbjct: 162 GTKEAKEKGLAQVAHLSKEEVAAVNEAKRQGRYTTD-------DGYIFSPTDIIDDLGDA 214
```

```
Query: 546 IIYPHGDHHHADPIDEHKP 564
            + PHG+H+H  P +   P
Sbjct: 215 YLVPHGNHYHYIPKKDLSP 233
```

---

tr   Q9ZHG7   **Hypothetical protein [Streptococcus agalactiae]** 822 AA
                                                                       align

```
 Score = 305 bits (781), Expect = 3e-81
 Identities = 198/612 (32%), Positives = 298/612 (48%), Gaps = 74/612 (12%)
```

```
Query: 1   MKFSKKYIAAGSAVIVSLSLCAYALNQHRSQENKDNNRVSYVDGSQSSQKSE--NLTPDQ 58
           MK +   YI + +A++++  + +Y L +H       +N+++Y+D S+  K+   N T DQ
Sbjct: 1   MKKTYGYIGSVAAILLATHIGSYQLGKHHMGLATKDNQIAYIDDSKGKVKAPKTNKTMDQ 60
```

```
Query: 59  VSQKEGIQAEQIVIKITDQGYVTSHGDHYHYYNGKVPYDALFSEELLMKDPNYQLKDADI 118
           +S +EGI AEQIV+KITDQGYVTSHGDHYH+YNGKVPYDA+ SEELLM DPNY  K +D+
Sbjct: 61  ISAEEGISAEQIVVKITDQGYVTSHGDHYHFYNGKVPYDAIISEELLMTDPNYHFKQSDV 120
```

```
Query: 119 VNEVKGGYIIKVDGKYYVYLKDAAHADNVRTKDEINRQKQEHVKD-NEKVNSNVA----- 172
           +NE+  GY+IKV+G YYVYLK  +   N+RTK +I  Q  + K+ EK  + VA
Sbjct: 121 INEILDGYVIKVNGNYYVYLKPGSKRKNIRTKQQIAEQVAKGTKEAKEKGLAQVAHLSKE 180
```

```
Query: 173 ------VARSQGRYTTNDGYVFNPADIIEDTGNAYIVPHGGHYHYIPXXXXXXXXXXXXX 226
                 A+ QGRYTT+DGY+F+P DII+D G+AY+VPHG HYHYIP
Sbjct: 181 EVAAVNEAKRQGRYTTDDGYIFSPTDIIDDLGDAYLVPHGNHYHYIPKKDLSPSELAAAQ 240
```

```
Query: 227 XXXXXXNMQPSQLS-YSSTASDNNTQSVAKGSTSKPAN-------------------- 263
                 + ++ S Y T +    ++    T P
Sbjct: 241 AYWSQKQGRGARPSDYRPTPAPGRRKAPIPDVTPNPGQGHQPDNGGYHPAPPRPNDASQN 300
```

```
Query: 264 -------KSENLQSLLKELYDSPSAQRYSESDGLVFDPAKIISRTPNGVAIPHGDHYHFI 316
                  K + + LL +L+      R+ E DGL+F+P ++I   G +PHGDHYH I
Sbjct: 301 KHQRDEFKGKTFKELLDQLHRLDLKYRHVEEDGLIFEPTQVIKSNAFGYVVPHGDHYHII 360
```

```
Query:  317  PYSKLSALEEKIARRVPISGT--GSTVSTNAKPN--EVVXXXXXXXXXXXXXXXXXKELSS  372
             P S+LS LE ++A R     T    + S ++KP+  EV                     +
Sbjct:  361  PRSQLSPLEMELADRYLAGQTDDNDSGSDHSKPSDKEVTHTFLGHRIKAYGKGLDGKPYD  420

Query:  373  ASDGYIFNPKDIVEETATAYIVRHGDHFHYIPKSNQIGQPTLPN--NSLATPSPSLPINP  430
             SD Y+F+ + I     +   +HGDHFHYI    ++ Q L    N+       +     +
Sbjct:  421  TSDAYVFSKESIHSVDKSGVTAKHGDHFHYI-GFGELEQYELDEVANWVKAKGQADELVA  479

Query:  431  GISHEKHEEDGYGFD----ANRIIAEDESGFIMSHGNHNHYFFKKDLTEEQIKAAQKHLE  486
             + E+ +E   FD   + ++ + + G+IM    ++++ + L    QI A++ L
Sbjct:  480  ALDQEQGKEKPL-FDTKKVSRKVTKDGKVGYIMPKDGKDYFYARYQLDLTQIAFAEQELM  538

Query:  487  EVKTSHNGLDSLSSHEQDYPGNAKEMKDLDKKIEEKIAGIMKQYGVKRESIVVNKEKNAI  546
                H  D                 +D  IE ++A +   +    +    +     +  +
Sbjct:  539  LKDKKHYRYDI----------------VDTGIEPRLAVDVSSLPMHAGNATYDTGSSFV  581

Query:  547  IYPHGDHHHADP  558
             I PH DH H  P
Sbjct:  582  I-PHIDHIHVVP  592


  Score = 47.8 bits (112), Expect = 0.001
  Identities = 53/313 (16%), Positives = 108/313 (33%), Gaps = 35/313 (11%)

Query:   67  AEQIVIKITDQGYVTSHGDHYHY--YNGKVPYD-----------------ALFSEELLM  106
             +++ + +   G   HGDH+HY +   Y+                   A  +E
Sbjct:  428  SKESIHSVDKSGVTAKHGDHFHYIGFGELEQYELDEVANWVKAKGQADELVAALDQEQGK  487

Query:  107  KDPNYQLKDAD--IVNEVKGGYIIKVDGKYYVYLKDAAHADNVRTKDEINRQKQEHVKDN  164
             + P + K     + + K GYI+ DGK Y Y +       +  ++      + +KD
Sbjct:  488  EKPLFDTKKVSRKVTKDGKVGYIMPKDGKDYFYARYQLDLTQIAFAEQ-----ELMLKDK  542

Query:  165  EKVNSNVAVARSQGRYTTN-DGYVFNPADIIEDTGNAYIVPHGGHYHYIPXXXXXXXXXX  223
             +    ++   + R +      + +  DTG++++++PH  H H +P
Sbjct:  543  KHYRYDIVDTGIEPRLAVDVSSLPMHAGNATYDTGSSFVIPHIDHIHVVPYSWLTRNQIA  602

Query:  224  XXXXXXXXXNMQPSQLSYSSTASDNNTQSVAKGSTSKPANKSENLQSLLKELYDSPSAQR  283
                      ++P  S        +       +   + +   +E+ + +   R
Sbjct:  603  TIKYVMQHPEVRPDVWSKPGHEESGSVIPNVTPLDKRAGMPNWQIIHSAEEVQKALAEGR  662

Query:  284  YSESDGLVFDPAKIISR-----TPNGVAIPH--GDHYHFIPYSKLSALEEKIARRVPISG  336
             ++ DG +FDP ++++         +IP  G    I S LS E + A+ +
Sbjct:  663  FAAPDGYIFDPRDVLAKETFVWKDGSFSIPRADGSSLRTINKSDLSQAEWQQAQELLAKK  722

Query:  337  TGSTVSTNAKPNE  349
              +   KP E
Sbjct:  723  NAGDATDTDKPEE  735


  Score = 38.1 bits (87), Expect = 0.88
  Identities = 39/199 (19%), Positives = 78/199 (38%), Gaps = 33/199 (16%)

Query:  372  SASDGYIFNPKDIVEETATAYIVRHGDHFHYIPKSNQIGQPTLPNNSLATPSPSLPINPG  431
             SA +G I  + +V+ T   Y+  HGDH+H+             +  +P +
Sbjct:   62  SAEEG-ISAEQIVVKITDQGYVTSHGDHYHFY------------------NGKVPYDAI  101

Query:  432  ISHEKHEED-GYGFDANRIIAEDESGFIMSHGNHNHYFFKKDLTEEQIKAAQKHLEEV--  488
             IS E   D  Y F  + +I E  G+++     + + K    + I+  Q+  E+V
```

```
Sbjct: 102  ISEELLMTDPNYHFKQSDVINEILDGYVIKVNGNYYVYLKPGSKRKNIRTKQQIAEQVAK 161

Query: 489  ---KTSHNGLDSLSSHEQDYPGNAKEMKDLDKKIEEKIAGIMKQYGVKRESIVVNKEKNA 545
               +     GL  ++    ++        E  K  +    +     +  +++    +A
Sbjct: 162  GTKEAKEKGLAQVAHLSKEEVAAVNEAKRQGRYTTD-------DGYIFSPTDIIDDLGDA 214

Query: 546  IIYPHGDHHHADPIDEHKP 564
              + PHG+H+H   P  + P
Sbjct: 215  YLVPHGNHYHYIPKKDLSP 233
```

```
tr Q99XV4  Hypothetical protein SPy2006 [SPy2006] [Streptococcus        825
           pyogenes]                                                    AA
                                                                        align
```

```
Score = 304 bits (779), Expect = 5e-81
Identities = 196/614 (31%), Positives = 297/614 (47%), Gaps = 76/614 (12%)

Query: 1    MKFSKKYIAAGSAVIVSLSLCAYALNQHRSQENKDNNRVSYVDGSQSSQKSE--NLTPDQ 58
            MK +  YI + +A++++  + +Y L +H        +N+++Y+D S+  K+   N T DQ
Sbjct: 1    MKKTYGYIGSVAAILLATHIGSYQLGKHHMGSATKDNQIAYIDDSKGKAKAPKTNKTMDQ 60

Query: 59   VSQKEGIQAEQIVIKITDQGYVTSHGDHYHYYNGKVPYDALFSEELLMKDPNYQLKDADI 118
            +S +EGI AEQIV+KITDQGYVTSHGDHYH+YNGKVPYDA+ SEELLM DPNY+ K +D+
Sbjct: 61   ISAEEGISAEQIVVKITDQGYVTSHGDHYHFYNGKVPYDAIISEELLMTDPNYRFKQSDV 120

Query: 119  VNEVKGGYIIKVDGKYYVYLKDAAHADNVRTKDEINRQKQEHVKD-NEKVNSNVA----- 172
            +NE+  GY+IKV+G YYVYLK     N+RTK +I  Q + K+ EK  + VA
Sbjct: 121  INEILDGYVIKVNGNYYVYLKPGSKRKNIRTKQQIAEQVAKGTKEAKEKGLAQVAHLSKE 180

Query: 173  ------VARSQGRYTTNDGYVFNPADIIEDTGNAYIVPHGGHYHYIP---XXXXXXXXXX 223
                  A+ QGRYTT+DGY+F+P DII+D G+AY+VPHG HYHYIP
Sbjct: 181  EVAAVNEAKRQGRYTTDDGYIFSPTDIIDDLGDAYLVPHGNHYHYIPKKDLSPSELAAAQ 240

Query: 224  XXXXXXXXXNMQPSQLSYSSTASDNNTQSVAKGSTSKPAN------------------- 263
                     +PS    +     ++      T  P
Sbjct: 241  AYWSQKQGRGARPSDYRPTPAPAPGRRKAPIPDVTPNPGQGHQPDNGGYHPAPPRPNDAS 300

Query: 264  ---------KSENLQSLLKELYDSPSAQRYSESDGLVFDPAKIISRTPNGVAIPHGDHYH 314
                     K +   + LL +L+      R+ E DGL+F+P ++I    G  +PHGDHYH
Sbjct: 301  QNKHQRDEFKGKTFKELLDQLHRLDLKYRHVEEDGLIFEPTQVIKSNAFGYVVPHGDHYH 360

Query: 315  FIPYSKLSALEEKIARRVPISGT--GSTVSTNAKPN--EVVXXXXXXXXXXXXXXXXKEL 370
            IP S+LS LE ++A R       T    + S ++KP+  EV                  +
Sbjct: 361  IIPRSQLSPLEMELADRYLAGQTEDDDSGSDHSKPSDKEVTHTFLGHRIKAYGKGLDGKP 420

Query: 371  SSASDGYIFNPKDIVEETATAYIVRHGDHFHYIPKSNQIGQPTLPN--NSLATPSPSLPI 428
            S  SD Y+F+ + I    +   +HGDHFHYI    ++  Q  L  N+    + +   +  +
Sbjct: 421  YDTSDAYVFSKESIHSVDKSGVTAKHGDHFHYI-GFGELEQYELDEVANWVKAKGQADEL 479

Query: 429  NPGISHEKHEEDGYGFD----ANRIIAEDESGFIMSHGNHNHYFFKKDLTEEQIKAAQKH 484
              + E+ +E   FD       + ++   + + G++M   ++++ +   L   QI A++
Sbjct: 480  AAALDQEQGKEKPL-FDTKKVSRKVTKDGKVGYMMPKDGKDYFYARDQLDLTQIAFAEQE 538

Query: 485  LEEVKTSHNGLDSLSSHEQDYPGNAKEMKDLDKKIEEKIAGIMKQYGVKRESIVVNKEKN 544
            L      H   D                +D   IE ++A +    +     +   + +  +
```

```
Sbjct:  539 LMLKDKKHYRYDI-----------------VDTGIEPRLAVDVSSLPMHAGNATYDTGSS 581

Query:  545 AIIYPHGDHHHADP 558
                +I PH DH H  P
Sbjct:  582 FVI-PHIDHIHVVP 594
```

```
 Score = 49.7 bits (117), Expect = 3e-04
 Identities = 53/313 (16%), Positives = 109/313 (33%), Gaps = 35/313 (11%)

Query:   67 AEQIVIKITDQGYVTSHGDHYHY--YNGKVPYD-----------------ALFSEELLM 106
               +++ +  +   G   HGDH+HY  +    Y+                 A  +E
Sbjct:  430 SKESIHSVDKSGVTAKHGDHFHYIGFGELEQYELDEVANWVKAKGQADELAAALDQEQGK 489

Query:  107 KDPNYQLKDAD--IVNEVKGGYIIKVDGKYYVYLKDAAHADNVRTKDEINRQKQEHVKDN 164
               + P + K    +   + K GY++  DGK Y Y +D    +  ++     +  +KD
Sbjct:  490 EKPLFDTKKVSRKVTKDGKVGYMMPKDGKDYFYARDQLDLTQIAFAEQ-----ELMLKDK 544

Query:  165 EKVNSNVAVARSQGRYTTN-DGYVFNPADIIEDTGNAYIVPHGGHYHYIPXXXXXXXXXX 223
               +    ++     + R +     + +    + +   DTG++++++PH  H H +P
Sbjct:  545 KHYRYDIVDTGIEPRLAVDVSSLPMHAGNATYDTGSSFVIPHIDHIHVVPYSWLTRDQIA 604

Query:  224 XXXXXXXXXNMQPSQLSYSSTASDNNTQSVAKGSTSKPANKSENLQSLLKELYDSPSAQR 283
                        ++P   S        +        +     +  +   +E+  + +   R
Sbjct:  605 TIKYVMQHPEVRPDIWSKPGHEESGSVIPNVTPLDKRAGMPNWQIIHSAEEVQKALAEGR 664

Query:  284 YSESDGLVFDPAKIISR-----TPNGVAIPH--GDHYHFIPYSKLSALEEKIARRVPISG 336
               ++  DG +FDP ++++          +IP  G    I  S LS  E + A+ +
Sbjct:  665 FATPDGYIFDPRDVLAKETFVWKDGSFSIPRADGSSLRTINKSDLSQAEWQQAQELLAKK 724

Query:  337 TGSTVSTNAKPNE 349
               +   KP E
Sbjct:  725 NAGDATDTDKPKE 737
```

```
 Score = 38.1 bits (87), Expect = 0.88
 Identities = 39/199 (19%), Positives = 78/199 (38%), Gaps = 33/199 (16%)

Query:  372 SASDGYIFNPKDIVEETATAYIVRHGDHFHYIPKSNQIGQPTLPNNSLATPSPSLPINPG 431
               SA +G I  +  +V+ T   Y+  HGDH+H+                  +  +P +
Sbjct:   62 SAEEG-ISAEQIVVKITDQGYVTSHGDHYHFY-------------------NGKVPYDAI 101

Query:  432 ISHEKHEED-GYGFDANRIIAEDESGFIMSHGNHNHYFFKKDLTEEQIKAAQKHLEEV-- 488
               IS E    D Y F + +I E  G+++     + + K    + I+   Q+  E+V
Sbjct:  102 ISEELLMTDPNYRFKQSDVINEILDGYVIKVNGNYYVYLKPGSKRKNIRTKQQIAEQVAK 161

Query:  489 ---KTSHNGLDSLSSHEQDYPGNAKEMKDLDKKIEEKIAGIMKQYGVKRESIVVNKEKNA 545
                  +   GL  ++   ++      E K  +  +          +   +  +++   +A
Sbjct:  162 GTKEAKEKGLAQVAHLSKEEVAAVNEAKRQGRYTTD-------DGYIFSPTDIIDDLGDA 214

Query:  546 IIYPHGDHHHADPIDEHKP 564
               + PHG+H+H  P +  P
Sbjct:  215 YLVPHGNHYHYIPKKDLSP 233
```

tr Q8K5Q1 **Histidine triad protein [SpyM3_1724] [Streptococcus pyogenes** 823 AA

(serotype M3)]   <u>align</u>

```
Score = 304 bits (778), Expect = 7e-81
Identities = 195/612 (31%), Positives = 295/612 (47%), Gaps = 74/612 (12%)

Query: 1    MKFSKKYIAAGSAVIVSLSLCAYALNQHRSQENKDNNRVSYVDGSQSSQKSE--NLTPDQ 58
            MK +  YI + +A++++  + +Y L +H        +N+++Y+D S+   K+   N T DQ
Sbjct: 1    MKKTYGYIGSVAAILLATHIGSYQLGKHHMGLATKDNQIAYIDDSKGKAKAPKTNKTMDQ 60

Query: 59   VSQKEGIQAEQIVIKITDQGYVTSHGDHYHYYNGKVPYDALFSEELLMKDPNYQLKDADI 118
            +S +EGI AEQIV+KITDQGYVTSHGDHYH+YNGKVPYDA+ SEELLM DPNY  K +D+
Sbjct: 61   ISAEEGISAEQIVVKITDQGYVTSHGDHYHFYNGKVPYDAIISEELLMTDPNYHFKQSDV 120

Query: 119  VNEVKGGYIIKVDGKYYVYLKDAAHADNVRTKDEINRQKQEHVKD-NEKVNSNVA----- 172
            +NE+  GY+IKV+G YYVY+K  +   N+RTK +I  Q  + K+  EK + VA
Sbjct: 121  INEILDGYVIKVNGNYYVYIKPGSKRKNIRTKQQIAEQVAKGTKEAKEKGLAQVAHLSKE 180

Query: 173  ------VARSQGRYTTNDGYVFNPADIIEDTGNAYIVPHGGHYHYIPXXXXXXXXXXXXX 226
                  A+ QGRYTT+DGY+F+P DII+D G+AY+VPHG HYHYIP
Sbjct: 181  EVAAVNEAKRQGRYTTDDGYIFSPTDIIDDLGDAYLVPHGNHYHYIPKKDLSPSELAAAQ 240

Query: 227  XXXXXXNMQPSQLS-YSSTASDNNTQSVAKGSTSKPAN--------------------- 263
                  + ++ S Y T +    ++    T  P
Sbjct: 241  AYWSQKQGRGARPSDYRPTPAPGRRKAPIPDVTPNPGQGHQPDNGGYHPAPPRPNDASQN 300

Query: 264  -------KSENLQSLLKELYDSPSAQRYSESDGLVFDPAKIISRTPNGVAIPHGDHYHFI 316
                   K +  + LL +L+      R+ E DGL+F+P ++I    G  +PHGDHYH I
Sbjct: 301  KHQRDEFKGKTFKELLDQLHRLDLKYRHVEEDGLIFEPTQVIKSNAFGYVVPHGDHYHII 360

Query: 317  PYSKLSALEEKIARRVPISGTGSTVS----TNAKPNEVVXXXXXXXXXXXXXXXXXKELSS 372
            P S+LS LE ++A R    T     S      + + EV                  +
Sbjct: 361  PRSQLSPLEMELADRYLAGQTEDNDSGSDHSKSSDKEVTHTFLGHRIKAYGKGLDGKPYD 420

Query: 373  ASDGYIFNPKDIVEETATAYIVRHGDHFHYIPKSNQIGQPTLPN--NSLATPSPSLPINP 430
            SD Y+F+ + I    +   ++HGDHFHYI    ++ Q L   N +     +   + +
Sbjct: 421  TSDAYVFSKESIHSVDKSGVTAKHGDHFHYI-GFGELEQYELDEVANWVKAKGQADELAA 479

Query: 431  GISHEKHEEDGYGFD----ANRIIAEDESGFIMSHGNHNHYFFKKDLTEEQIKAAQKHLE 486
            +  E+ +E      FD   + ++  + + G+IM     +++  ++++ +  L  QI A++ L
Sbjct: 480  ALDQEQGKEKPL-FDTKKVSRKVTKDGKVGYIMPKDGKDYFYARDQLDLTQIAFAEQELM 538

Query: 487  EVKTSHNGLDSLSSHEQDYPGNAKEMKDLDKKIEEKIAGIMKQYGVKRESIVVNKEKNAI 546
            +H  D                +D IE ++A +  +   +   + +   + +
Sbjct: 539  LKDKNHYRYDI----------------VDTGIEPRLAVDVSSLPMHAGNATYDTGSSFV 581

Query: 547  IYPHGDHHHADP 558
            I PH DH H  P
Sbjct: 582  I-PHIDHIHVVP 592


Score = 50.8 bits (120), Expect = 1e-04
Identities = 54/313 (17%), Positives = 108/313 (34%), Gaps = 35/313 (11%)

Query: 67   AEQIVIKITDQGYVTSHGDHYHY--YNGKVPYD-----------------ALFSEELLM 106
            +++ +  +   G   HGDH+HY   Y+         A  +E
Sbjct: 428  SKESIHSVDKSGVTAKHGDHFHYIGFGELEQYELDEVANWVKAKGQADELAAALDQEQGK 487
```

```
Query:  107  KDPNYQLKDAD--IVNEVKGGYIIKVDGKYYVYLKDAAHADNVRTKDEINRQKQEHVKDN  164
             + P +  K     + + K GYI+ DGK Y Y +D      +   ++      +  +KD
Sbjct:  488  EKPLFDTKKVSRKVTKDGKVGYIMPKDGKDYFYARDQLDLTQIAFAEQ-----ELMLKDK  542


Query:  165  EKVNSNVAVARSQGRYTTN-DGYVFNPADIIEDTGNAYIVPHGGHYHYIPXXXXXXXXXX  223
                 ++    + + R  +     +  +   DTG++++PH  H H +P
Sbjct:  543  NHYRYDIVDTGIEPRLAVDVSSLPMHAGNATYDTGSSFVIPHIDHIHVVPYSWLTRDQIA  602


Query:  224  XXXXXXXXXNMQPSQLSYSSTASDNNTQSVAKGSTSKPANKSENLQSLLKELYDSPSAQR  283
                      ++P  S       +        +   + + +    +E+ + +  R
Sbjct:  603  TIKYVMQHPEVRPDVWSKPGHEESGSVIPNVTPLDKRAGMPNWQIIHSAEEVQKALAEGR  662


Query:  284  YSESDGLVFDPAKIISR-----TPNGVAIPH--GDHYHFIPYSKLSALEEKIARRVPISG  336
             ++  DG +FDP ++++        +IP  G     I  S LS  E + A+ +
Sbjct:  663  FATPDGYIFDPRDVLAKETFVWKDGSFSIPRADGSSLRTINKSDLSQAEWQQAQELLAKK  722


Query:  337  TGSTVSTNAKPNE  349
                 +   KP E
Sbjct:  723  NAGDATDTDKPKE  735
```

```
  Score = 38.1 bits (87), Expect = 0.88
  Identities = 39/199 (19%), Positives = 78/199 (38%), Gaps = 33/199 (16%)

Query:  372  SASDGYIFNPKDIVEETATAYIVRHGDHFHYIPKSNQIGQPTLPNNSLATPSPSLPINPG  431
             SA +G I  + +V+ T   Y+ HGDH+H+                     +  +P +
Sbjct:   62  SAEEG-ISAEQIVVKITDQGYVTSHGDHYHFY------------------NGKVPYDAI  101


Query:  432  ISHEKHEED-GYGFDANRIIAEDESGFIMSHGNHNHYFFKKDLTEEQIKAAQKHLEEV--  488
             IS E    D Y F + +I E  G+++  + + + K    + I+ Q+  E+V
Sbjct:  102  ISEELLMTDPNYHFKQSDVINEILDGYVIKVNGNYYVYIKPGSKRKNIRTKQQIAEQVAK  161


Query:  489  ---KTSHNGLDSLSSHEQDYPGNAKEMKDLDKKIEEKIAGIMKQYGVKRESIVVNKEKNA  545
                +   GL ++    ++      E K  +  +     +    + +++    +A
Sbjct:  162  GTKEAKEKGLAQVAHLSKEEVAAVNEAKRQGRYTTD-------DGYIFSPTDIIDDLGDA  214


Query:  546  IIYPHGDHHHADPIDEHKP  564
              + PHG+H+H  P +  P
Sbjct:  215  YLVPHGNHYHYIPKKDLSP  233
```

```
  tr   Q93GT5      Histidine triad protein of group A streptococci      825 AA
                   [htpA]                                                align
                   [Streptococcus pyogenes]


  Score =  302 bits (774), Expect = 2e-80
  Identities = 195/614 (31%), Positives = 297/614 (47%), Gaps = 76/614 (12%)

Query:    1  MKFSKKYIAAGSAVIVSLSLCAYALNQHRSQENKDNNRVSYVDGSQSSQKSE--NLTPDQ   58
             MK +  YI + +A++++  + +Y L +H        +N+++Y+D S+  K+   N T DQ
Sbjct:    1  MKKTYGYIGSVAAILLATHIGSYQLGKHHMGSATKDNQIAYIDDSKGKAKAPKTNKTMDQ   60


Query:   59  VSQKEGIQAEQIVIKITDQGYVTSHGDHYHYYNGKVPYDALFSEELLMKDPNYQLKDADI  118
             +S +EGI AEQIV+KITDQGYVTSHGDHYH+YNGKVPYDA+ SEELLM DPNY+ K +D+
Sbjct:   61  ISAEEGISAEQIVVKITDQGYVTSHGDHYHFYNGKVPYDAIISEELLMTDPNYRFKQSDV  120
```

```
Query: 119 VNEVKGGYIIKVDGKYYVYLKDAAHADNVRTKDEINRQKQEHVKD-NEKVNSNVA----- 172
            +NE+  GY+IKV+G YYVYLK +   N+RTK +I Q + K+  EK + VA
Sbjct: 121 INEILDGYVIKVNGNYYVYLKPGSKRKNIRTKQQIAEQVAKGTKEAKEKGLAQVAHLSKE 180

Query: 173 ------VARSQGRYTTNDGYVFNPADIIEDTGNAYIVPHGGHYHYIP---XXXXXXXXXX 223
                  A+ QGRYTT+DGY+F+P DII+D G+AY+VPHG HYHYIP
Sbjct: 181 EVAAVNEAKRQGRYTTDDGYIFSPTDIIDDLGDAYLVPHGNHYHYIPKKDLSPSELAAAQ 240

Query: 224 XXXXXXXXXNMQPSQLSYSSTASDNNTQSVAKGSTSKPAN------------------- 263
                     +PS   +    ++    T   P
Sbjct: 241 AYWSQKQGRGARPSDYRPTPAPAPGRRKAPIPDVTPNPGQGHQPDNGGYHPAPPRPNDAS 300

Query: 264 ---------KSENLQSLLKELYDSPSAQRYSESDGLVFDPAKIISRTPNGVAIPHGDHYH 314
                     K + + LL +L+ +    R+ E DGL+F+P ++I     G +PHGDH+H
Sbjct: 301 QNKHQRDEFKGKTFKELLDQLHRLDLKYRHVEEDGLIFEPTQVIKSNAFGYVVPHGDHHH 360

Query: 315 FIPYSKLSALEEKIARRVPISGT--GSTVSTNAKPN--EVVXXXXXXXXXXXXXXXXXXXKEL 370
            IP S+LS LE ++A R    T   + S ++KP+ EV                        +
Sbjct: 361 IIPRSQLSPLEMELADRYLAGQTEDDDSGSDHSKPSDKEVTHTFLGHRIKAYGKGLDGKP 420

Query: 371 SSASDGYIFNPKDIVEETATAYIVRHGDHFHYIPKSNQIGQPTLPN--NSLATPSPSLPI 428
            SD Y+F+ + I  +    +HGDHFHYI   ++ Q L  N+   + +
Sbjct: 421 YDTSDAYVFSKESIHSVDKSGVTAKHGDHFHYI-GFGELEQYELDEVANWVKAKGQADEL 479

Query: 429 NPGISHEKHEEDGYGFD----ANRIIAEDESGFIMSHGNHNHYFFKKDLTEEQIKAAQKH 484
            +  E+ +E    FD  + ++ + + G++M   ++++ + L  QI A++
Sbjct: 480 AAALDQEQGKEKPL-FDTKKVSRKVTKDGKVGYMMPKDGKDYFYARDQLDLTQIAFAEQE 538

Query: 485 LEEVKTSHNGLDSLSSHEQDYPGNAKEMKDLDKKIEEKIAGIMKQYGVKRESIVVNKEKN 544
            L   H   D                +D  IE ++A +   +   +   +     +
Sbjct: 539 LMLKDKKHYRYDI-----------------VDTGIEPRLAVDVSSLPMHAGNATYDTGSS 581

Query: 545 AIIYPHGDHHHADP 558
            +I PH DH H  P
Sbjct: 582 FVI-PHIDHIHVVP 594
```

```
Score = 49.7 bits (117), Expect = 3e-04
Identities = 53/313 (16%), Positives = 109/313 (33%), Gaps = 35/313 (11%)

Query: 67  AEQIVIKITDQGYVTSHGDHYHY--YNGKVPYD-----------------ALFSEELLM 106
            +++ + + G   HGDH+HY +   Y+                    A  +E
Sbjct: 430 SKESIHSVDKSGVTAKHGDHFHYIGFGELEQYELDEVANWVKAKGQADELAAALDQEQGK 489

Query: 107 KDPNYQLKDAD--IVNEVKGGYIIKVDGKYYVYLKDAAHADNVRTKDEINRQKQEHVKDN 164
            + P + K   + + K GY++  DGK Y Y +D  +     + ++   + +KD
Sbjct: 490 EKPLFDTKKVSRKVTKDGKVGYMMPKDGKDYFYARDQLDLTQIAFAEQ-----ELMLKDK 544

Query: 165 EKVNSNVAVARSQGRYTTN-DGYVFNPADIIEDTGNAYIVPHGGHYHYIPXXXXXXXXXX 223
            +    ++    + R +    + +      DTG++++++PH  H H +P
Sbjct: 545 KHYRYDIVDTGIEPRLAVDVSSLPMHAGNATYDTGSSFVIPHIDHIHVVPYSWLTRDQIA 604

Query: 224 XXXXXXXXXNMQPSQLSYSSTASDNNTQSVAKGSTSKPANKSENLQSLLKELYDSPSAQR 283
            ++P   S     +       +   + +  +E+ + +   R
Sbjct: 605 TIKYVMQHPEVRPDIWSKPGHEESGSVIPNVTPLDKRAGMPNWQIIHSAEEVQKALAEGR 664

Query: 284 YSESDGLVFDPAKIISR-----TPNGVAIPH--GDHYHFIPYSKLSALEEKIARRVPISG 336
            ++  DG +FDP ++++       +IP G     I S LS E + A+ +
```

```
Sbjct: 665 FATPDGYIFDPRDVLAKETFVWKDGSFSIPRADGSSLRTINKSDLSQAEWQQAQELLAKK 724

Query: 337 TGSTVSTNAKPNE 349
                +   KP E
Sbjct: 725 NAGDATDTDKPKE 737
```

```
 Score = 38.1 bits (87), Expect = 0.88
 Identities = 39/199 (19%), Positives = 78/199 (38%), Gaps = 33/199 (16%)

Query: 372 SASDGYIFNPKDIVEETATAYIVRHGDHFHYIPKSNQIGQPTLPNNSLATPSPSLPINPG 431
            SA +G I   + +V+ T   Y+  HGDH+H+                   +    +P +
Sbjct: 62  SAEEG-ISAEQIVVKITDQGYVTSHGDHYHFY------------------NGKVPYDAI 101

Query: 432 ISHEKHEED-GYGFDANRIIAEDESGFIMSHGNHNHYFFKKDLTEEQIKAAQKHLEEV-- 488
            IS E    D  Y F + +I E   G+++     + + + K    + I+  Q+  E+V
Sbjct: 102 ISEELLMTDPNYRFKQSDVINEILDGYVIKVNGNYYVYLKPGSKRKNIRTKQQIAEQVAK 161

Query: 489 ---KTSHNGLDSLSSHEQDYPGNAKEMKDLDKKIEEKIAGIMKQYGVKRESIVVNKEKNA 545
               +    GL ++   ++       E K  +   +      +    + +++    +A
Sbjct: 162 GTKEAKEKGLAQVAHLSKEEVAAVNEAKRQGRYTTD-------DGYIFSPTDIIDDLGDA 214

Query: 546 IIYPHGDHHHADPIDEHKP 564
            +  PHG+H+H  P  +  P
Sbjct: 215 YLVPHGNHYHYIPKKDLSP 233
```

```
tr Q8E338  Hypothetical protein gbs1925 [gbs1925] [Streptococcus       481
           agalactiae                                                   AA
           (serotype III)]                                              align
```

```
 Score =  301 bits (771), Expect = 4e-80
 Identities = 163/380 (42%), Positives = 221/380 (57%), Gaps = 57/380 (15%)

Query: 8   IAAGSAVIVSLSLCAYALNQHRSQENKDNNRVSYVDGSQSSQKSE----NLTPDQVSQKE 63
            I+A S +   + AY L + +Q K +N VSYV   +S K++    N TPDQ+S++E
Sbjct: 9   ISALSVALFGTGVGAYQLGSYNAQ--KSDNSVSYVKTDKSDSKAQATAVNKTPDQISKEE 66

Query: 64  GIQAEQIVIKITDQGYVTSHGDHYHYYNGKVPYDALFSEELLMKDPNYQLKDADIVNEVK 123
            GI AEQIV+KITD GYVTSHGDHYHYYNGKVPYDA+ SEEL+MKDP+Y   AD++NEVK
Sbjct: 67  GISAEQIVVKITDDGYVTSHGDHYHYYNGKVPYDAIISEELIMKDPSYVFNKADVINEVK 126

Query: 124 GGYIIKVDGKYYVYLKDAAHADNVRTKDEINRQKQEHVKDNEK-------------VNSN 170
            GYIIKV+GKYY+YLK+ +    NVRTK++I +Q++E K  K                ++++
Sbjct: 127 DGYIIKVNGKYYLYLKEGSKRTNVRTKEQIQKQREEWSKGGSKGESGKHSSAKTQALSAS 186

Query: 171 VAVARSQGRYTTNDGYVFNPADIIEDTGNAYIVPHGGHYHYIPXXXXXXXXXXXXX----- 225
            V A++ GRYTT+DGYVF+P D+I+D G+A++VPHG H+HYIP
Sbjct: 187 VREAKASGRYTTDDGYVFSPTDVIDDMGDAFLVPHGDHFHYIPKADLSPSELSAAQAYWN 246

Query: 226 -XXXXXXXNMQPSQLS----------------------------YSSTASDNNTQS 252
             + +PS  S                                    +S A ++ T
Sbjct: 247 RKTGRSGNSSKPSNSSSYIHASAPSGNVSTGRHANAPISIPRVTHANHWSKPAGNHATAP 306

Query: 253 VAKGSTSKPANKSENLQSLLKELYDSPSAQRYSESDGLVFDPAKIISRTPNGVAIPHGDH 312
            T+KP NK   L +LK LY P   R+ ESDGLV+DPA++ + T  GV+IPHG+H
```

```
Sbjct: 307 KHHAPTTKPINKDSALDKMLKRLYAQPLYARHVESDGLVYDPAQVNAFTAIGVSIPHGNH 366

Query: 313 YHFIPYSKLSALEEKIARRV 332
           +HFI Y  +S LE +  R V
Sbjct: 367 FHFIHYKDMSPLELEATRMV 386
```

```
 Score = 74.7 bits (182), Expect = 9e-12
 Identities = 74/304 (24%), Positives = 124/304 (40%), Gaps = 40/304 (13%)
```

```
Query: 186 YVFNPADIIEDTGNAYIVPHGGHYHYIPXXXXXXXXXXXXXXXXXXXXNMQPSQLSYSSTA 245
           YVFN AD+I +  + YI+  G Y+                        +Q  +   +S
Sbjct: 114 YVFNKADVINEVKDGYIIKVNGKYYLY------LKEGSKRTNVRTKEQIQKQREEWSKGG 167

Query: 246 SDNNTQSVAKGSTSKPANKSENLQSLLKELYDSPSAQRYSESDGLVFDPAKIISRTPNGV 305
           S          KG + K + S  Q+L   + ++ ++ RY+  DG VF P +I     +
Sbjct: 168 S--------KGESGK--HSSAKTQALSASVREAKASGRYTTDDGYVFSPTDVIDDMGDAF 217

Query: 306 AIPHGDHYHFIPYSKLSALEEKIARRVPISGTGSTVSTNAKPNEVVXXXXXXXXXXXXXX 365
           +PHGDH+H+IP + LS  E   A+       TG +   ++KP+
Sbjct: 218 LVPHGDHFHYIPKADLSPSELSAAQAYWNRKTGRS-GNSSKPSN----------SSSYIH 266

Query: 366 XXKELSSASDGYIFN-PKDIVEET-ATAYIVRHGDHF----HYIPKSNQIGQPTLPNNSL 419
             + S G    N P I   TA +    G+H    H+ P + I +     +++L
Sbjct: 267 ASAPSGNVSTGRHANAPISIPRVTHANHWSKPAGNHATAPKHHAPTTKPINK----DSAL 322

Query: 420 ATPSPSLPINPGISHEKH-EEDGYGFDANRIIAEDESGFIMSHGNHNHYFFKKDLTEEQI 478
           L  P  + +H E DG +D ++ A    G  + HGNH H+    KD++  ++
Sbjct: 323 DKMLKRLYAQP--LYARHVESDGLVYDPAQVNAFTAIGVSIPHGNHFHFIHYKDMSPLEL 380

Query: 479 KAAQ 482
           +A +
Sbjct: 381 EATR 384
```

```
 Score = 57.8 bits (138), Expect = 1e-06
 Identities = 59/293 (20%), Positives = 112/293 (38%), Gaps = 83/293 (28%)
```

```
Query: 240 SYSSTASDNNTQSVAKGSTSKPANKSENLQSLLKELYDSPSAQRYSESDGLVFDPAKIIS 299
           SY++ SDN+   V          K++  S +   + + +  S+ +G+ +     ++
Sbjct: 28  SYNAQKSDNSVSYV---------KTDKSDSKAQATAVNKTPDQISKEEGISAEQI-VVK 76

Query: 300 RTPNGVAIPHGDHYHF----IPYSKLSALEEKIARRVPISGTGSTVSTNAKPNEVVXXXX 355
           T +G   HGDHYH+     +PY + +
Sbjct: 77  ITDDGYVTSHGDHYHYYNGKVPYDAIIS----------------------------- 104

Query: 356 XXXXXXXXXXXXXKELSSASDGYIFNPKDIVEETATAYIVR-HGDHFHYIPKSN------- 407
                        +EL      Y+FN  D++  E     YI++ +G ++ Y+ + +
Sbjct: 105 -----------EELIMKDPSYVFNKADVINEVKDGYIIKVNGKYYLYLKEGSKRTNVRT 152

Query: 408 --QIGQP------------TLPNNSLATPSPSLPINPGISHEKHE-EDGYGFDANRIIAE 452
           QI +               + ++S T + S +    + ++ +DGY F    +I +
Sbjct: 153 KEQIQKQREEWSKGGSKGESGKHSSAKTQALSASVREAKASGRYTTDDGYVFSPTDVIDD 212

Query: 453 DESGFIMSHGNHNHYFFKKDLTEEQIKAAQKHLEEVKTSHNGLDSLSSHEQDY 505
           F++ HG+H HY K DL+  ++ AAQ +   KT +G  S S+  Y
Sbjct: 213 MGDAFLVPHGDHFHYIPKADLSPSELSAAQAYWNR-KTGRSGNSSKPSNSSSY 264
```

```
Score = 41.6 bits (96), Expect = 0.080
Identities = 54/216 (25%), Positives = 84/216 (38%), Gaps = 39/216 (18%)

Query: 384  IVEETATAYIVRHGDHFHYIPKSNQIGQPTLPNNSLATPSPSLPINPGISHEKHEED-GY 442
            +V+ T   Y+  HGDH+HY                   +  +P +  IS E  +D  Y
Sbjct: 74   VVKITDDGYVTSHGDHYHYY------------------NGKVPYDAIISEELIMKDPSY  114

Query: 443  GFDANRIIAEDESGFIMSHGNHNHYFFKKD-------LTEEQIKAAQKHLEEVKTSHNGL 495
            F+   +I E + G+I+   N +Y + K+       T+EQI  QK  EE   S  G
Sbjct: 115  VFNKADVINEVKDGYIIK-VNGKYYLYLKEGSKRTNVRTKEQI---QKQREE--WSKGG-  167

Query: 496  DSLSSHEQDYPGNAKEMKDLDKKIEEKIAG--IMKQYGVKRESIVVNKEKNAIIYPHGDH  553
               S  E   +AK     E K +G    V  + V++   +A +  PHGDH
Sbjct: 168  ---SKGESGKHSSAKTQALSASVREAKASGRYTTDDGYVFSPTDVIDDMGDAFLVPHGDH  224

Query: 554  HHADPIDEHKPVGIGHSHSNYELFKPEEGVAKKEGN 589
            H  P + P +   + + +        G + K N
Sbjct: 225  FHYIPKADLSPSELSAAQAYWNRKTGRSGNSSKPSN 260
```

tr <u>Q877Y2</u> **Hypothetical protein SPs1722 [SPs1722] [Streptococcus pyogenes** 794 AA
(serotype M3)]

<u>align</u>

```
Score =  291 bits (746), Expect = 3e-77
Identities = 187/577 (32%), Positives = 277/577 (47%), Gaps = 74/577 (12%)

Query: 36   NNRVSYVDGSQSSQKSE--NLTPDQVSQKEGIQAEQIVIKITDQGYVTSHGDHYHYYNGK 93
            +N+++Y+D S+   K+   N T DQ+S +EGI AEQIV+KITDQGYVTSHGDHYH+YNGK
Sbjct: 7    DNQIAYIDDSKGKAKAPKTNKTMDQISAEEGISAEQIVVKITDQGYVTSHGDHYHFYNGK  66

Query: 94   VPYDALFSEELLMKDPNYQLKDADIVNEVKGGYIIKVDGKYYVYLKDAAHADNVRTKDEI 153
            VPYDA+ SEELLM DPNY  K +D++NE+  GY+IKV+G YYVY+K   +   N+RTK +I
Sbjct: 67   VPYDAIISEELLMTDPNYHFKQSDVINEILDGYVIKVNGNYYVYIKPGSKRKNIRTKQQI  126

Query: 154  NRQKQEHVKD-NEKVNSNVA----------VARSQGRYTTNDGYVFNPADIIEDTGNAY 201
             Q + K+ EK + VA           A+ QGRYTT+DGY+F+P DII+D G+AY
Sbjct: 127  AEQVAKGTKEAKEKGLAQVAHLSKEEVAAVNEAKRQGRYTTDDGYIFSPTDIIDDLGDAY  186

Query: 202  IVPHGGHYHYIPXXXXXXXXXXXXXXXXXXXXXXNMQPSQLS-YSSTASDNNTQSVAKGSTSK 260
            +VPHG HYHYIP                  +  ++ S Y  T +     ++     T
Sbjct: 187  LVPHGNHYHYIPKKDLSPSELAAAQAYWSQKQGRGARPSDYRPTPAPGRRKAPIPDVTPN  246

Query: 261  PAN-------------------------KSENLQSLLKELYDSPSAQRYSESDGLV 291
            P                         K +  + LL +L+      R+ E DGL+
Sbjct: 247  PGQGHQPDNGGYHPAPPRPNDASQNKHQRDEFKGKTFKELLDQLHRLDLKYRHVEEDGLI  306

Query: 292  FDPAKIISRTPNGVAIPHGDHYHFIPYSKLSALEEKIARRVPISGTGSTVS----TNAKP 347
            F+P ++I    G +PHGDHYH IP S+LS LE ++A R   T    S   + +
Sbjct: 307  FEPTQVIKSNAFGYVVPHGDHYHIIPRSQLSPLEMELADRYLAGQTEDNDSGSDHSKSSD  366

Query: 348  NEVVXXXXXXXXXXXXXXXXXXKELSSASDGYIFNPKDIVEETATAYIVRHGDHFHYIPKSN 407
            EV                  +    SD Y+F+ + I     +   +HGDHFHYI
Sbjct: 367  KEVTHTFLGHRIKAYGKGLDGKPYDTSDAYVFSKESIHSVDKSGVTAKHGDHFHYI-GFG  425
```

```
Query:  408  QIGQPTLPN--NSLATPSPSLPINPGISHEKHEEDGYGFD----ANRIIAEDESGFIMSH  461
             ++ Q   L    N+    +     +      E+ +E   FD     + ++   + + G+IM
Sbjct:  426  ELEQYELDEVANWVKAKGQADELAAALDQEQGKEKPL-FDTKKVSRKVTKDGKVGYIMPK  484

Query:  462  GNHNHYFFKKDLTEEQIKAAQKHLEEVKTSHNGLDSLSSHEQDYPGNAKEMKDLDKKIEE  521
             ++++ +  L   QI A++ L    +H   D                  +D   IE
Sbjct:  485  DGKDYFYARDQLDLTQIAFAEQELMLKDKNHYRYDI-----------------VDTGIEP  527

Query:  522  KIAGIMKQYGVKRESIVVNKEKNAIIYPHGDHHHADP  558
             ++A +   +    +    +     + +I PH DH H  P
Sbjct:  528  RLAVDVSSLPMHAGNATYDTGSSFVI-PHIDHIHVVP  563


 Score = 50.8 bits (120), Expect = 1e-04
 Identities = 54/313 (17%), Positives = 108/313 (34%), Gaps = 35/313 (11%)

Query:   67  AEQIVIKITDQGYVTSHGDHYHY--YNGKVPYD-----------------ALFSEELLM  106
             +++ +  +    G     HGDH+HY  +     Y+                  A   +E
Sbjct:  399  SKESIHSVDKSGVTAKHGDHFHYIGFGELEQYELDEVANWVKAKGQADELAAALDQEQGK  458

Query:  107  KDPNYQLKDAD--IVNEVKGGYIIKVDGKYYVYLKDAAHADNVRTKDEINRQKQEHVKDN  164
             + P + K   +   + K GYI+ DGK Y Y +D     +    ++    +  +KD
Sbjct:  459  EKPLFDTKKVSRKVTKDGKVGYIMPKDGKDYFYARDQLDLTQIAFAEQ-----ELMLKDK  513

Query:  165  EKVNSNVAVARSQGRYTTN-DGYVFNPADIIEDTGNAYIVPHGGHYHYIPXXXXXXXXXX  223
             ++     + R +    + +     DTG+++++PH  H H +P
Sbjct:  514  NHYRYDIVDTGIEPRLAVDVSSLPMHAGNATYDTGSSFVIPHIDHIHVVPYSWLTRDQIA  573

Query:  224  XXXXXXXXXNMQPSQLSYSSTASDNNTQSVAKGSTSKPANKSENLQSLLKELYDSPSAQR  283
             ++P   S     +     +        + + +       +E+   + +  R
Sbjct:  574  TIKYVMQHPEVRPDVWSKPGHEESGSVIPNVTPLDKRAGMPNWQIIHSAEEVQKALAEGR  633

Query:  284  YSESDGLVFDPAKIISR-----TPNGVAIPH--GDHYHFIPYSKLSALEEKIARRVPISG  336
             ++  DG +FDP  ++++          +IP  G    I  S LS  E + A+ +
Sbjct:  634  FATPDGYIFDPRDVLAKETFVWKDGSFSIPRADGSSLRTINKSDLSQAEWQQAQELLAKK  693

Query:  337  TGSTVSTNAKPNE  349
             +    KP E
Sbjct:  694  NAGDATDTDKPKE  706


 Score = 38.1 bits (87), Expect = 0.88
 Identities = 39/199 (19%), Positives = 78/199 (38%), Gaps = 33/199 (16%)

Query:  372  SASDGYIFNPKDIVEETATAYIVRHGDHFHYIPKSNQIGQPTLPNNSLATPSPSLPINPG  431
             SA +G I  + +V+ T   Y+ HGDH+H+                  +   +P +
Sbjct:   33  SAEEG-ISAEQIVVKITDQGYVTSHGDHYHFY-----------------NGKVPYDAI   72

Query:  432  ISHEKHEED-GYGFDANRIIAEDESGFIMSHGNHNHYFFKKDLTEEQIKAAQKHLEEV--  488
             IS E   D Y F + +I E  G+++   + + + K   + I+  Q+  E+V
Sbjct:   73  ISEELLMTDPNYHFKQSDVINEILDGYVIKVNGNYYVYIKPGSKRKNIRTKQQIAEQVAK  132

Query:  489  ---KTSHNGLDSLSSHEQDYPGNAKEMKDLDKKIEEKIAGIMKQYGVKRESIVVNKEKNA  545
                +    GL  ++   ++      E K  +    +      + + +++    +A
Sbjct:  133  GTKEAKEKGLAQVAHLSKEEVAAVNEAKRQGRYTTD-------DGYIFSPTDIIDDLGDA  185

Query:  546  IIYPHGDHHHADPIDEHKP  564
             + PHG+H+H  P  + P
```

```
Sbjct: 186 YLVPHGNHYHYIPKKDLSP 204
```

tr Q9AE21 **Hypothetical protein (Fragment) [Streptococcus agalactiae]** 289 AA

align

```
 Score = 241 bits (614), Expect = 7e-62
 Identities = 117/227 (51%), Positives = 159/227 (69%), Gaps = 14/227 (6%)

Query: 1   MKFSKKYIAAGSAVIVSLSLCAYALNQHRSQENKDNNRVSYVDGSQSSQKSE--NLTPDQ 58
            MK +   YI + +A++++  + +Y L +H        +N+++Y+D S+   K+    N T DQ
Sbjct: 1   MKKTYGYIGSVAAILLATHIGSYQLGKHHMGLATKDNQIAYIDDSKGKVKAPKTNKTMDQ 60

Query: 59  VSQKEGIQAEQIVIKITDQGYVTSHGDHYHYYNGKVPYDALFSEELLMKDPNYQLKDADI 118
            +S +EGI AEQIV+KITDQGYVTSHGDHYH+YNGKVPYDA+ SEELLM DPNY+ K +D+
Sbjct: 61  ISAEEGISAEQIVVKITDQGYVTSHGDHYHFYNGKVPYDAIISEELLMTDPNYRFKQSDV 120

Query: 119 VNEVKGGYIIKVDGKYYVYLKDAAHADNVRTKDEINRQKQEHVKD-NEKVNSNVA----- 172
            +NE+  GY+IKV+G YYVYLK +   N+RTK +I Q +  K+ EK  + VA
Sbjct: 121 INEILDGYVIKVNGNYYVYLKPGSKRKNIRTKQQIAEQVAKGTKEAKEKGLAQVAHLSKE 180

Query: 173 ------VARSQGRYTTNDGYVFNPADIIEDTGNAYIVPHGGHYHYIP 213
                  A+ QGRYTT+DGY+F+P DII+D G+AY+VPHG HYHYIP
Sbjct: 181 EVAAVNEAKRQGRYTTDDGYIFSPTDIIDDLGDAYLVPHGNHYHYIP 227
```

```
 Score = 56.2 bits (134), Expect = 3e-06
 Identities = 33/94 (35%), Positives = 48/94 (50%), Gaps = 24/94 (25%)

Query: 372 SASDGYIFNPKDIVEETATAYIVRHGDHFHYIPK--------------SNQIGQPTLPN 416
            +  DGYIF+P DI+++   AY+V HG+H+HYIPK             S + G+   P+
Sbjct: 195 TTDDGYIFSPTDIIDDLGDAYLVPHGNHYHYIPKKDLSPSELAAAQAYWSQKQGRGARPS 254

Query: 417 NSLATPSP---SLPI-----NPGISHEKHEEDGY 442
            +  TP+P    PI     NPG H+ + GY
Sbjct: 255 DYRPTPAPGRRKAPIPDVTPNPGQGHQP-DNGGY 287
```

```
 Score = 49.7 bits (117), Expect = 3e-04
 Identities = 45/214 (21%), Positives = 74/214 (34%), Gaps = 70/214 (32%)

Query: 297 IISRTPNGVAIPHGDHYHF----IPYSKLSALEEKIARRVPISGTGSTVSTNAKPNEVVX 352
            ++  T G   HGDHYHF     +PY + +
Sbjct: 73  VVKITDQGYVTSHGDHYHFYNGKVPYDAIIS---------------------------- 103

Query: 353 XXXXXXXXXXXXXXXKELSSASDGYIFNPKDIVEETATAYIVR-HGDHFHYIP-----KS 406
                           +EL   Y F   D++ E    Y+++ +G+++ Y+      K+
Sbjct: 104 --------------EELLMTDPNYRFKQSDVINEILDGYVIKVNGNYYVYLKPGSKRKN 148

Query: 407 NQIGQPTLPNNSLATPSPSLPINPGISHEKHEE---------------DGYGFDANRII 450
            +  Q    + T         ++H EE               DGY F   II
Sbjct: 149 IRTKQQIAEQVAKGTKEAKEKGLAQVAHLSKEEVAAVNEAKRQGRYTTDDGYIFSPTDII 208

Query: 451 AEDESGFIMSHGNHNHYFFKKDLTEEQIKAAQKH 484
            +    +++ HGNH HY  KKDL+ ++ AAQ +
```

Sbjct: 209 DDLGDAYLVPHGNHYHYIPKKDLSPSELAAAQAY 242

 Score = 38.1 bits (87), Expect = 0.88
 Identities = 39/199 (19%), Positives = 78/199 (38%), Gaps = 33/199 (16%)

Query: 372 SASDGYIFNPKDIVEETATAYIVRHGDHFHYIPKSNQIGQPTLPNNSLATPSPSLPINPG 431
            SA +G I   + +V+ T   Y+ HGDH+H+   +   +P +
Sbjct: 62 SAEEG-ISAEQIVVKITDQGYVTSHGDHYHFY------------------NGKVPYDAI 101

Query: 432 ISHEKHEED-GYGFDANRIIAEDESGFIMSHGNHNHYFFKKDLTEEQIKAAQKHLEEV-- 488
            IS E   D Y F + +I E G+++   + + + K   + I+   Q+   E+V
Sbjct: 102 ISEELLMTDPNYRFKQSDVINEILDGYVIKVNGNYYVYLKPGSKRKNIRTKQQIAEQVAK 161

Query: 489 ---KTSHNGLDSLSSHEQDYPGNAKEMKDLDKKIEEKIAGIMKQYGVKRESIVVNKEKNA 545
               +   GL ++   ++   E K   +   +   +   + +++   +A
Sbjct: 162 GTKEAKEKGLAQVAHLSKEEVAAVNEAKRQGRYTTD-------DGYIFSPTDIIDDLGDA 214

Query: 546 IIYPHGDHHHADPIDEHKP 564
             + PHG+H+H   P + P
Sbjct: 215 YLVPHGNHYHYIPKKDLSP 233

tr   Q8DQ06   **Pneumococcal histidine triad protein E, truncation   182 AA**
                   [phtE-truncation] [Streptococcus pneumoniae (strain   align
                   ATCC
                   BAA-255 / R6)]

 Score = 121 bits (304), Expect = 6e-26
 Identities = 67/156 (42%), Positives = 91/156 (57%), Gaps = 16/156 (10%)

Query: 370 LSSASDGYIFNPKDIVEETATAYIVRHGDHFHYIPKSNQIGQPTLPN-NSLATPSPSLPI 428
            +S+ DGY+FNP DIV +T AYIVRHGDH+HYIPKS+   P+ N   + + S S+
Sbjct: 9   VSAVDDGYVFNPNDIVRDTGDAYIVRHGDHYHYIPKSSLNNPPSHSNTEEVGSSSSSSVLS 68

Query: 429 NPGISHEKHEEDGYGFDANRIIAEDESGFIMSHGNHNHYFFKKDLTEEQIKAAQKHLEEV 488
            NP +   EEDG+GFDANRII+ED GF++ HG+HNHY   IK   K   E
Sbjct: 69 NPSLHVHHEEEDGHGFDANRIISEDSEGFVIPHGDHNHY----------IKVQTKGYE-- 116

Query: 489 KTSHNGLDSLSSHEQDYPGNAKEMKDLDKKIEEKIA 524
               N + SL S   +YP   + K + K+++ +A
Sbjct: 117 AALKNKIPSLQS---NYPPGTFDEKAVLAKVDQLLA 149

 Score = 87.4 bits (215), Expect = 1e-15
 Identities = 54/161 (33%), Positives = 74/161 (45%), Gaps = 39/161 (24%)

Query: 173 VARSQGRYTTNDGYVFNPADIIEDTGNAYIVPHGGHYHYIPXXXXXXXXXXXXXXXXXXXX 232
            V + G   +DGYVFNP DI+ DTG+AYIV HG HYHYIP
Sbjct: 3   VTPNNGVSAVDDGYVFNPNDIVRDTGDAYIVRHGDHYHYIP------------------ 43

Query: 233 NMQPSQLSYSSTASDNNTQSVAKGSTSKPANKSENLQSLLKELYDSPSAQRYSESDGLVF 292
            + S ++ S +NT+ V S+S +N S ++   + E DG F
Sbjct: 44 -----KSSLNNPPSHSNTEEVGSSSSSVLSNPSLHV-------------HHEEEDGHGF 84

Query: 293 DPAKIISRTPNGVAIPHGDHYHFIPYSKLSALEEKIARRVP 333

```
                   D  +IIS     G  IPHGDH H+I   +    E  +  ++P
Sbjct:  85  DANRIISEDSEGFVIPHGDHNHYIKV-QTKGYEAALKNKIP  124


 Score = 46.6 bits (109), Expect = 0.002
 Identities = 32/135 (23%), Positives = 58/135 (42%), Gaps = 34/135 (25%)

Query: 426  LPINP--GISHEKHEEDGYGFDANRIIAEDESGFIMSHGNHNHYFFKKDLTEEQIKAAQK  483
            +P+ P  G+S       +DGY F+ N I+ +    +I+ HG+H HY  K  L       +
Sbjct:   1  MPVTPNNGVSAV---DDGYVFNPNDIVRDTGDAYIVRHGDHYHYIPKSSLNNPP---SHS   54

Query: 484  HLEEVKTSHNGL---DSLSSHEQDYPGNAKEMKDLDKKIEEKIAGIMKQYGVKRESIVVN  540
            + EEV +S + +    SL  H ++   G+  +                         +++
Sbjct:  55  NTEEVGSSSSSVLSNPSLHVHHEEEDGHGFDANR--------------------IIS   91

Query: 541  KEKNAIIYPHGDHHH  555
            ++     + PHGDH+H
Sbjct:  92  EDSEGFVIPHGDHNH  106




tr Q8E029  Hypothetical protein SAG0907 [SAG0907] [Streptococcus     877
           agalactiae                                                AA
           (serotype V)]                                             align


 Score =  111 bits (277), Expect = 8e-23
 Identities = 109/406 (26%), Positives = 170/406 (41%), Gaps = 59/406 (14%)

Query: 250  TQSVAKGSTSKPANKSENLQSLLKELYDSPSAQRYSESDG-LVFDPAKIISRTPNGVAIP  308
            T+ VAK   K ANK +    +   +  +   +   DG L+  +KI+S T +G+ +
Sbjct:  36  TKQVAK----KKANKKKVSVKESHKKRKGVAGVDFPTDDGFLLTKDSKILSHTDSGIVVA   91

Query: 309  HGDHYHFIPYSKLSALEEKIARRVPISGTGSTVSTNAKPNEVVXXXXXXXXXXXXXXXXXXK  368
            HG+H HFI YS L       K + +P         +T     N+ V
Sbjct:  92  HGNHSHFIFYSDLKG--SKFSYLIP------NANTKTNKNQAVRNFKAGAVA--------  135

Query: 369  ELSSASDGYIFNPKDIVEETATAYIVRHGDHFHYIPKSNQIGQPTLPNNSLATPSPSLPI  428
            +++ +DGY+FNP DIV E A Y+VRHGDHFHYIPK++  Q +      A+ + S
Sbjct: 136  -VNTLNDGYVFNPADIVSEDANGYVVRHGDHFHYIPKASLSQQKQVQ----ASRAVSRLG  190

Query: 429  NPGISHEK----------HEEDGYGFDANRIIAEDESGFIMSHGNHNHYFFKKDLTEEQ  477
            N   SH +              +G+ F+  I    +G ++ H NH H+    DL +
Sbjct: 191  NQSNSHYRVNSSKIAGLHYPTSNGFLFNGRGIKGTTPTGILVEHHNHLHFISFADLRKGG  250

Query: 478  IKAAQKHLEEVKTSHNGLDSLSSHEQDYPGNAKEMKDLDKKIEEKIAGIMKQYGVKRESI  537
            +      +  K + +    S SS +       +    L K I++K+A + ++  +      I
Sbjct: 251  WGSIADRYQPQKKADSKKQSPSSKK------PRTENTLPKDIKDKLAYLARELHLDISRI  304

Query: 538  VVNKEKNAII---YPHGDHHH---ADPIDEHKPVGIGH--------SHSNYELFKPEEGV  583
            V K N I    YPH DH H   A  ID KP+     H          H + E     E
Sbjct: 305  RVLKTLNGEIGFEYPHDDHTHVIMAKDIDLSKPIPNPHHDDEDHHKGHHHDESDHKHEEH  364

Query: 584  AKKEGNKVYTGEELTNVVNLLKNSTFN-NQNFTLANGQKRVSFSFP  628
            + NK+ + E+  ++ L +   N NQ  L +     + F +P
Sbjct: 365  EHTKSNKL-SDEDQKKLIYLAEKLGLNPNQIEVLTSEDGSIIFKYP  409
```

```
 Score =  100 bits (250), Expect = 1e-19
 Identities = 107/468 (22%), Positives = 192/468 (40%), Gaps = 92/468 (19%)

Query: 139 KDAAHADNVRTKDEINRQKQEHVKDNEKVNSNVAVARSQGRYTTNDGYVFNP-ADIIEDT 197
            K++ H    +   +   +K+ VK++ K   VA      + T+DG++   + I+  T
Sbjct: 29  KNSHHIKTKQVAKKKANKKKVSVKESHKKRKGVAGVD----FPTDDGFLLTKDSKILSHT 84

Query: 198 GNAYIVPHGGHYHYIPXXXXXXXXXXXXXXXXXXXXXXNMQPSQLSYSSTASDNNTQSVAKGS 257
            +  +V HG H H+I                         +++ S+ SY    ++  T
Sbjct: 85  DSGIVVAHGNHSHFI----------------FYSDLKGSKFSYLIPNANTKT------- 120

Query: 258 TSKPANKSENLQSLLKELYDSPSAQRYSESDGLVFDPAKIISRTPNGVAIPHGDHYHFIP 317
               NK++ +++     + + +    + +DG VF+PA I+S   NG  + HGDH+H+IP
Sbjct: 121 -----NKNQAVRN-----FKAGAVAVNTLNDGYVFNPADIVSEDANGYVVRHGDHFHYIP 170

Query: 318 YSKLSALEEKIARRVPISGTGSTVSTNAKPNEVVXXXXXXXXXXXXXXXXXXXKELSSASDGY 377
             + LS  ++  A R  +S  G+  +++ + N                          S+G+
Sbjct: 171 KASLSQQKQVQASRA-VSRLGNQSNSHYRVNSSKIAGLHY-------------PTSNGF 215

Query: 378 IFNPKDIVEETATAYIVRHGDHFHYI-----------------PKSNQIGQPTLPNNSL 419
            +FN + I   T T  +V H +H H+I                 P+     +   P++
Sbjct: 216 LFNGRGIKGTTPTGILVEHHNHLHFISFADLRKGGWGSIADRYQPQKKADSKKQSPSSKK 275

Query: 420 ATPSPSLP--INPGISHEKHEEDGYGFDANRI----IAEDESGFIMSHGNHNHYFFKKDL 473
            +LP   I    +++     E       D +RI          E GF     H +H H   KD+
Sbjct: 276 PRTENTLPKDIKDKLAYLARE---LHLDISRIRVLKTLNGEIGFEYPHDDHTHVIMAKDI 332

Query: 474 T-EEQIKAAQKHLEEVKTSHNGLDSLSSHEQDYPGNAKEMKDLDKKIEEKIAGIMKQYGV 532
            + I     E+    H+   +S   HE+    + ++ D D+K    K+   + ++ G+
Sbjct: 333 DLSKPIPNPHHDDEDHHKGHHHDESDHKHEEHEHTKSNKLSDEDQK---KLIYLAEKLGL 389

Query: 533 KRESI-VVNKEKNAII--YPHGDHHH---ADPIDEHKPVGIG-HSHSN 573
             I V+  E  +II  YPH DH H   +   I+  KP+  G H HS+
Sbjct: 390 NPNQIEVLTSEDGSIIFKYPHDDHSHTIASKDIEIGKPIPDGHHDHSH 437


 Score = 52.8 bits (125), Expect = 3e-05
 Identities = 70/346 (20%), Positives = 120/346 (34%), Gaps = 84/346 (24%)

Query: 21  CAYALNQHRSQENKDNNRVSY-----------VDGSQSSQKSENLTPDQVSQKEG--IQ 66
            C +AL   +QE+K+++ +          V    +S +K + +          +G  +
Sbjct: 16  CHFALTACHTQEHKNSHHIKTKQVAKKKANKKKVSVKESHKKRKGVAGVDFPTDDGFLLT 75

Query: 67  AEQIVIKITDQGYVTSHGDHYHYYNGKVPYDALFSEELLMKDPN--------------- 110
            +  ++ TD G V +HG+H H+     + Y L  +        PN
Sbjct: 76  KDSKILSHTDSGIVVAHGNHSHF----IFYSDLKGSKFSYLIPNANTKTNKNQAVRNFKA 131

Query: 111 ----------YQLKDADIVNEVKGGYIIKVDGKYYVYLKDAAHADNV-----RTKDEIN 154
                      Y    ADIV+E   GY+++   G  ++ Y+  A+ +       R   +
Sbjct: 132 GAVAVNTLNDGYVFNPADIVSEDANGYVVR-HGDHFHYIPKASLSQQKQVQASRAVSRLG 190

Query: 155 RQKQEHVKDNEKVNSNVAVARSQGRYTTNDGYVFNPADIIEDTGNAYIVPHGGHYHYIPX 214
            Q   H + N    + +         Y T++G++FN  I   T    +V H  H H+I
Sbjct: 191 NQSNSHYRVNSSKIAGL-------HYPTSNGFLFNGRGIKGTTPTGILVEHHNHLHFISF 243

Query: 215 XXXXXXXXXXXXXXXXXXXXNMQPSQLSYSSTASDNNTQSVAKGSTSKPANKSENLQSLLKE 274
                                 + S A    Q  A     P++K   ++ L +
Sbjct: 244 ADLRKG-----------------GWGSIADRYQPQKKADSKKQSPSSKKPRTENTLPK 284
```

```
Query: 275 LYDSPSAQRYSESDGLVFDPAKI-ISRTPN---GVAIPHGDHYHFI 316
                      A    E   L  D ++I + +T N   G   PH DH H I
Sbjct: 285 DIKDKLAYLARE---LHLDISRIRVLKTLNGEIGFEYPHDDHTHVI 327
```

```
tr  Q8E5R2  Hypothetical protein gbs0918 [gbs0918] [Streptococcus      877
            agalactiae                                                  AA
            (serotype III)]                                             align
```

```
 Score =  109 bits (273), Expect = 2e-22
 Identities = 110/406 (27%), Positives = 169/406 (41%), Gaps = 59/406 (14%)

Query: 250 TQSVAKGSTSKPANKSENLQSLLKELYDSPSAQRYSESDG-LVFDPAKIISRTPNGVAIP 308
            T+ VAK    K ANK +     +     +  +  DG L+    +KI+S   +G+ +
Sbjct: 36  TKQVAK----KKANKKKVSVKESHKKRKGVAGVDFPTDDGFLLTKDSKILSHPDSGIVVA 91

Query: 309 HGDHYHFIPYSKLSALEEKIARRVPISGTGSTVSTNAKPNEVVXXXXXXXXXXXXXXXXXXK 368
            HG+H HFI YS L       K + +P        + NAK N+
Sbjct: 92  HGNHSHFIFYSDLKG--SKFSYLIP--------NANAKTNKKQAVRNFKAGAVA------ 135

Query: 369 ELSSASDGYIFNPKDIVEETATAYIVRHGDHFHYIPKSNQIGQPTLPNNSLATPSPSLPI 428
            +++ +DGY+FNP DIV E A  Y+VRHGDHFHYIPK++   Q +    A+ + S
Sbjct: 136 -VNTLNDGYVFNPADIVSEDANGYVVRHGDHFHYIPKASLSQQKQVQ----ASRAVSRLG 190

Query: 429 NPGISHEK----------HEEDGYGFDANRIIAEDESGFIMSHGNHNHYYFFKKDLTEEQ 477
            N   SH +              DG+ F+   I    +G ++ H NH H+    DL +
Sbjct: 191 NQNNSHYRVNSSKIAGLHHPTSDGFLFNGQGIKGTTPTGILVEHHNHLHFISFADLRKGG 250

Query: 478 IKAAQKHLEEVKTSHNGLDSLSSHEQDYPGNAKEMKDLDKKIEEKIAGIMKQYGVKRESI 537
            +     + K   +  S SS +       +    L K I++K+A + ++   +     I
Sbjct: 251 WGSIADRYQPQKKVDSKKQSPSSKK------PRTENTLPKDIKDKLAYLARELHLDISRI 304

Query: 538 VVNKEKNAII---YPHGDHHH---ADPIDEHKPVGIGH--------SHSNYELFKPEEGV 583
            V K N I   YPH DH H   A  ID KP+  H         H + E       E
Sbjct: 305 RVLKTLNGEIGFEYPHDDHTHVIMAKDIDLSKPIPNPHHDDEDHHKGHHHDESDHKHEEH 364

Query: 584 AKKEGNKVYTGEELTNVVNLLKNSTFN-NQNFTLANGQKRVSFSFP 628
            + NK+ + E+   ++ L +   N NQ  L +     + F +P
Sbjct: 365 EHTKSNKL-SDEDQKKLIYLAEKLGLNPNQIEVLTSEDGSIIFKYP 409
```

```
 Score =  100 bits (248), Expect = 2e-19
 Identities = 110/471 (23%), Positives = 186/471 (39%), Gaps = 98/471 (20%)

Query: 139 KDAAHADNVRTKDEINRQKQEHVKDNEKVNSNVAVARSQGRYTTNDGYVFNP-ADIIEDT 197
            K++ H   +   + +K+ VK++ K  VA       + T+DG++   + I+
Sbjct: 29  KNSHHIKTKQVAKKKANKKKVSVKESHKKRKGVAGVD----FPTDDGFLLTKDSKILSHP 84

Query: 198 GNAYIVPHGGHYHYIPXXXXXXXXXXXXXXXXXXXXXNMQPSQLSY---SSTASDNNTQSVA 254
            +  +V HG H H+I                         +++ S+ SY    ++ A  N  Q+V
Sbjct: 85  DSGIVVAHGNHSHFI----------------FYSDLKGSKFSYLIPNANAKTNKKQAVR 127

Query: 255 KGSTSKPANKSENLQSLLKELYDSPSAQRYSESDGLVFDPAKIISRTPNGVAIPHGDHYH 314
                 A + N                DG VF+PA I+S   NG  + HGDH+H
Sbjct: 128 NFKAGAVAVNTLN------------------DGYVFNPADIVSEDANGYVVRHGDHFH 167
```

```
Query:  315  FIPYSKLSALEEKIARRVPISGTGSTVSTNAKPNEVVXXXXXXXXXXXXXXXXXXKELSSAS 374
             +IP + LS  ++   A R  +S  G+   +++ + N                          S
Sbjct:  168  YIPKASLSQQKQVQASRA-VSRLGNQNNSHYRVNSSKIAGLHH-------------PTS 212

Query:  375  DGYIFNPKDIVEETATAYIVRHGDHFHYI----------------PKSNQIGQPTLPN 416
             DG++FN + I   T T  +V H +H H+I                P+     +   P+
Sbjct:  213  DGFLFNGQGIKGTTPTGILVEHHNHLHFISFADLRKGGWGSIADRYQPQKKVDSKKQSPS 272

Query:  417  NSLATPSPSLP--INPGISHEKHEEDGYGFDANRI----IAEDESGFIMSHGNHNHYFFK 470
             +       +LP  I  +++   E      D +RI       E GF   H +H H
Sbjct:  273  SKKPRTENTLPKDIKDKLAYLARE---LHLDISRIRVLKTLNGEIGFEYPHDDHTHVIMA 329

Query:  471  KDLT-EEQIKAAQKHLEEVKTSHNGLDSLSSHEQDYPGNAKEMKDLDKKIEEKIAGIMKQ 529
             KD+  + I     E+    H+  +S   HE+   + ++ D D+K  K+  + ++
Sbjct:  330  KDIDLSKPIPNPHHDDEDHHKGHHHDESDHKHEEHEHTKSNKLSDEDQK---KLIYLAEK 386

Query:  530  YGVKRESI-VVNKEKNAII--YPHGDHHH---ADPIDEHKPVGIG-HSHSN 573
             G+    I V+  E  +II  YPH DH H    +  I+ KP+  G H HS+
Sbjct:  387  LGLNPNQIEVLTSEDGSIIFKYPHDDHSHTIASKDIEIGKPIPDGHHDHSH 437
```

 Score = 48.9 bits (115), Expect = 5e-04
 Identities = 68/345 (19%), Positives = 120/345 (34%), Gaps = 82/345 (23%)

```
Query:   21  CAYALNQHRSQENKDNNRVSY-----------VDGSQSSQKSENLTPDQVSQKEG--IQ 66
             C +AL   +QE+K+++ +              V   +S +K + +          +G  +
Sbjct:   16  CHFALTACHTQEHKNSHHIKTKQVAKKKANKKKVSVKESHKKRKGVAGVDFPTDDGFLLT 75

Query:   67  AEQIVIKITDQGYVTSHGDHYHYYNGKVPYDALFSEELLMKDPN--------------- 110
             + + ++  D G V +HG+H H+   + Y  L  +        PN
Sbjct:   76  KDSKILSHPDSGIVVAHGNHSHF----IFYSDLKGSKFSYLIPNANAKTNKKQAVRNFKA 131

Query:  111  -----------YQLKDADIVNEVKGGYIIKVDGKYYVYLKDAAHADNV-----RTKDEIN 154
                       Y    ADIV+E   GY+++  G ++ Y+  A+  +       R    +
Sbjct:  132  GAVAVNTLNDGYVFNPADIVSEDANGYVVR-HGDHFHYIPKASLSQQKQVQASRAVSRLG 190

Query:  155  RQKQEHVKDNEKVNSNVAVARSQGRYTTNDGYVFNPADIIEDTGNAYIVPHGGHYHYIPX 214
              Q   H + N    + +     +    + T+DG++FN  I   T    +V H   H H+I
Sbjct:  191  NQNNSHYRVNSSKIAGL-------HHPTSDGFLFNGQGIKGTTPTGILVEHHNHLHFI-- 241

Query:  215  XXXXXXXXXXXXXXXXXXXXNMQPSQLSYSSTASDNNTQSVAKGSTSKPANKSENLQSLLKE 274
                                 QP +   S  S ++ +    +   + P + + L  L +E
Sbjct:  242  ---SFADLRKGGWGSIADRYQPQKKVDSKKQSPSSKK--PRTENTLPKDIKDKLAYLARE 296

Query:  275  LYDSPSAQRYSESDGLVFDPAKIISRTPN---GVAIPHGDHYHFI 316
             L+   S  R       + +T N   G   PH DH H I
Sbjct:  297  LHLDISRIR-------------VLKTLNGEIGFEYPHDDHTHVI 327
```

tr Q8P0G5 **Putative internalin A [spyM18_1373] [Streptococcus pyogenes** 792 AA
            (serotype M18)]
                                                                        align

 Score = 99.4 bits (246), Expect = 3e-19
 Identities = 93/350 (26%), Positives = 143/350 (40%), Gaps = 72/350 (20%)

```
Query:  237  SQLSYSSTASDNN------TQSVAKGSTS---KPANKSENLQSLLKELYDSPSAQRYSES  287
             SQL+  +  S  N       T+   KG TS    KP  KS+ +    K +   +    +
Sbjct:   16  SQLTLIACQSRGNGTYPIKTKQSRKGMTSNKIKPIKKSKKINKTHKGV----AGVDFPTD   71

Query:  288  DGLVFDP-AKIISRTPNGVAIPHGDHYHFIPYSKL--SALEEKIARRVPISGTGSTVSTN  344
             DG +    +KI+S+T G+ + H  H HFI Y+  L   S  E   I +     +
Sbjct:   72  DGFILTKDSKILSKTDQGIVVDHDGHSHFIFYADLKGSPFEYLIPKGASL----------  121

Query:  345  AKPNEVVXXXXXXXXXXXXXXXXXXXKELSSASDGYIFNPKDIVEETATAYIVRHGDHFHYIP  404
             AKP                       +++      Y FNP DIV E A  Y VRH DHFHYI
Sbjct:  122  AKP---------AVAQQAASQGTSKVADPHHHYEFNPADIVAEDALGYTVRHDDHFHYIL  172

Query:  405  KSNQIGQPTLPNNSLAT--PSPSLPIN-------PGISHEKHEEDGYGFDANRIIAEDES  455
             KS+  GQ    +AT  P  S P++        PG+       DG+ F+   I+   ++
Sbjct:  173  KSSLSGQTQAQAKQVATRLPQTSSPVSTVTANGIPGLHFP--TSDGFQFNGQGIVGVTKA  230

Query:  456  GFIMSHGNHNHYFFKKDLTE-------EQIKAAQKHLEEVKTSHNGLDSLSSHEQDYPGN  508
             +++H  H H      DL +        +Q   A+K  +   +T H
Sbjct:  231  SILVNHDGHLHPISFADLRQGGWAHVADQYDPAKKAEKPAETHH----------------  274

Query:  509  AKEMKDLDKKIEEKIAGIMKQYGVKRESI---VVNKEKNAIIYPHGDHHH  555
             E+ +  +K+ +EK+A + ++ G+    +I        K  + YPH DH H
Sbjct:  275  TPELSEREKEYQEKLAYLAEKLGIDPSTIKRVETQDGKLGLEYPHHDHAH  324


   Score = 51.6 bits (122), Expect = 8e-05
   Identities = 73/360 (20%), Positives = 135/360 (37%), Gaps = 68/360 (18%)

Query:    1  MKFSKKYIAAGSAVIVSLSLCA--------YALNQHRSQENKDNNRVSYVDGSQSSQKS-   51
             MK   K   I  G  +   L+L A         Y +   +S++   +N++   +   S+   K+
Sbjct:    1  MKTKKVIILVGLLLLSSQLTLIACQSRGNGTYPIKTKQSRKGMTSNKIKPIKKSKKINKTH   60

Query:   52  ENLTPDQVSQKEG--IQAEQIVIKITDQGYVTSHGDHYHY----------YNGKVPYDAL   99
             + +            +G  +   +  ++ TDQG V  H  H H+              +   +P  A
Sbjct:   61  KGVAGVDFPTDDGFILTKDSKILSKTDQGIVVDHDGHSHFIFYADLKGSPFEYLIPKGAS  120

Query:  100  FSEELLMK-----------DPN--YQLKDADIVNEVKGGYIIKVDGKYYVYLKDAAHADN  146
             ++  + +           DP+   Y+    ADIV E   GY ++ D   ++  LK +
Sbjct:  121  LAKPAVAQQAASQGTSKVADPHHHYEFNPADIVAEDALGYTVRHDDHFHYILKSSLSG--  178

Query:  147  VRTKDEINRQKQEHVKDNEKVNSNVAVARSQG----RYTTNDGYVFNPADIIEDTGNAYI  202
             +     Q ++     + +S V+    + G        + T+DG+ FN   I+   T + +
Sbjct:  179  -----QTQAQAKQVATRLPQTSSPVSTVTANGIPGLHFPTSDGFQFNGQGIVGVTKASIL  233

Query:  203  VPHGGHYHYIPXXXXXXXXXXXXXXXXXXXXXNMQPSQLSYSSTASDNNTQSVAKGSTSKPA  262
             V H GH H I                       P++ +   A  ++T +++      +
Sbjct:  234  VNHDGHLHPI-----SFADLRQGGWAHVADQYDPAKKA-EKPAETHHTPELSE----REK  283

Query:  263  NKSENLQSLLKELYDSPSA-QRYSESDGLVFDPAKIISRTPNGVAIPHGDHYHFIPYSKL  321
             +  E L  L ++L  PS  +R    DG +            G+  PH DH H +   S +
Sbjct:  284  EYQEKLAYLAEKLGIDPSTIKRVETQDGKL-----------GLEYPHHDHAHVLMLSDI  331
```

tr Q8K714 **Putative internalin A [inlA] [Streptococcus pyogenes (serotype** 792
   M3)]   AA

```
 Score = 97.4 bits (241), Expect = 1e-18
 Identities = 93/350 (26%), Positives = 140/350 (39%), Gaps = 72/350 (20%)

Query: 237  SQLSYSSTASDNN------TQSVAKGSTS---KPANKSENLQSLLKELYDSPSAQRYSES 287
            SQL+ +  S N       T+   KG TS   KP   KS+    K +     +      +
Sbjct: 16   SQLTLIACQSRGNGTYPIKTKQSRKGMTSNKIKPIKKSKKTNKTHKGV----AGVDFPTD 71

Query: 288  DGLVFDP-AKIISRTPNGVAIPHGDHYHFIPYSKL--SALEEKIARRVPISGTGSTVSTN 344
            DG +    +KI+S+T G+ + H H HFI Y+ L   S E  I +   +
Sbjct: 72   DGFILTKDSKILSKTDQGIVVDHDGHSHFIFYADLKGSPFEYLIPKGASL---------- 121

Query: 345  AKPNEVVXXXXXXXXXXXXXXXXXXXKELSSASDGYIFNPKDIVEETATAYIVRHGDHFHYIP 404
            AKP                       +++     Y FNP DIV E A  Y VRH DHFHYI
Sbjct: 122  AKP---------AVAQQAASQGTSKVADPHHHYEFNPADIVAEDALGYTVRHDDHFHYIL 172

Query: 405  KSNQIGQPTLPNNSLAT--PSPSLPIN-------PGISHEKHEEDGYGFDANRIIAEDES 455
            KS+  GQ    +AT  P  S P++        PG+       DG+ F+   I+    +
Sbjct: 173  KSSLSGQTQAQAKQVATRLPQTSSPVSTATANGIPGLHFP--TSDGFQFNGQGIVGVTKD 230

Query: 456  GFIMSHGNHNHYFFKKDLTE-------EQIKAAQKHLEEVKTSHNGLDSLSSHEQDYPGN 508
            ++ H  H H      DL +         +Q  A+K   +  +T H
Sbjct: 231  NILVDHDGHLHPISFADLRQGGWAHVADQYDPAKKAEKPAETHH--------------- 274

Query: 509  AKEMKDLDKKIEEKIAGIMKQYGVKRESI---VVNKEKNAIIYPHGDHHH 555
            E+ + +K+ +EK+A + ++ G+   +I       K   YPH DH H
Sbjct: 275  TPELSEREKEYQEKLAYLAEKLGIDPSTIKRVETQDGKLGLEYPHHDHAH 324
```

```
 Score = 54.7 bits (130), Expect = 9e-06
 Identities = 74/360 (20%), Positives = 137/360 (37%), Gaps = 68/360 (18%)

Query: 1    MKFSKKYIAAGSAVIVSLSLCA--------YALNQHRSQENKDNNRVSYVDGSQSSQKS- 51
            MK  K  I  G  +  L+L A        Y +   +S++  +N++ +   S+ + K+
Sbjct: 1    MKTKKVIILVGLLLSSQLTLIACQSRGNGTYPIKTKQSRKGMTSNKIKPIKKSKKTNKTH 60

Query: 52   ENLTPDQVSQKEG--IQAEQIVIKITDQGYVTSHGDHYHY----------YNGKVPYDAL 99
            + +          +G  +  +  ++ TDQG V  H  H H+         +    +P  A
Sbjct: 61   KGVAGVDFPTDDGFILTKDSKILSKTDQGIVVDHDGHSHFIFYADLKGSPFEYLIPKGAS 120

Query: 100  FSEELLMK----------DPN--YQLKDADIVNEVKGGYIIKVDGKYYVYLKDAAHADN 146
            ++  + +           DP+  Y+   ADIV E   GY ++ D  ++ LK +
Sbjct: 121  LAKPAVAQQAASQGTSKVADPHHHYEFNPADIVAEDALGYTVRHDDHFHYILKSSLSG-- 178

Query: 147  VRTKDEINRQKQEHVKDNEKVNSNVAVARSQG----RYTTNDGYVFNPADIIEDTGNAYI 202
            +   Q ++        + +S V+ A + G        + T+DG+ FN  I+  T +  +
Sbjct: 179  -----QTQAQAKQVATRLPQTSSPVSTATANGIPGLHFPTSDGFQFNGQGIVGVTKDNIL 233

Query: 203  VPHGGHYHYIPXXXXXXXXXXXXXXXXXXXXXXNMQPSQLSYSSTASDNNTQSVAKGSTSKPA 262
            V H GH H I                       P++ +    A  ++T +++    +
Sbjct: 234  VDHDGHLHPI-----SFADLRQGGWAHVADQYDPAKKA-EKPAETHHTPELSE----REK 283

Query: 263  NKSENLQSLLKELYDSPSA-QRYSESDGLVFDPAKIISRTPNGVAIPHGDHYHFIPYSKL 321
             E L  L ++L  PS +R     DG +        G+  PH DH H +   S +
Sbjct: 284  EYQEKLAYLAEKLGIDPSTIKRVETQDGKL-----------GLEYPHHDHAHVLMLSDI 331
```

tr <u>Q79XH7</u> **Putative internalin A [SPs0825] [Streptococcus pyogenes**   792
               (serotype   AA
               M3)]   <u>align</u>

Score = 97.4 bits (241), Expect = 1e-18
Identities = 93/350 (26%), Positives = 140/350 (39%), Gaps = 72/350 (20%)

Query: 237 SQLSYSSTASDNN------TQSVAKGSTS---KPANKSENLQSLLKELYDSPSAQRYSES 287
            SQL+ + S N   T+   KG TS   KP KS+   K +   +   +
Sbjct: 16   SQLTLIACQSRGNGTYPIKTKQSRKGMTSNKIKPIKKSKKTNKTHKGV----AGVDFPTD 71

Query: 288 DGLVFDP-AKIISRTPNGVAIPHGDHYHFIPYSKL--SALEEKIARRVPISGTGSTVSTN 344
            DG +   +KI+S+T G+ + H H HFI Y+ L   S E I +   +
Sbjct: 72   DGFILTKDSKILSKTDQGIVVDHDGHSHFIFYADLKGSPFEYLIPKGASL---------- 121

Query: 345 AKPNEVVXXXXXXXXXXXXXXXXXXKELSSASDGYIFNPKDIVEETATAYIVRHGDHFHYIP 404
            AKP   +++   Y FNP DIV E A Y VRH DHFHYI
Sbjct: 122 AKP---------AVAQQAASQGTSKVADPHHHYEFNPADIVAEDALGYTVRHDDHFHYIL 172

Query: 405 KSNQIGQPTLPNNSLAT--PSPSLPIN-------PGISHEKHEEDGYGFDANRIIAEDES 455
            KS+ GQ   +AT P S P++   PG+   DG+ F+   I+   +
Sbjct: 173 KSSLSGQTQAQAKQVATRLPQTSSPVSTATANGIPGLHFP--TSDGFQFNGQGIVGVTKD 230

Query: 456 GFIMSHGNHNHYFFKKDLTE-------EQIKAAQKHLEEVKTSHNGLDSLSSHEQDYPGN 508
            ++ H H H   DL +   +Q A+K   + +T H
Sbjct: 231 NILVDHDGHLHPISFADLRQGGWAHVADQYDPAKKAEKPAETHH--------------- 274

Query: 509 AKEMKDLDKKIEEKIAGIMKQYGVKRESI---VVNKEKNAIIYPHGDHHH 555
            E+ + +K+ +EK+A + ++ G+   +I   K + YPH DH H
Sbjct: 275 TPELSEREKEYQEKLAYLAEKLGIDPSTIKRVETQDGKLGLEYPHHDHAH 324

Score = 54.7 bits (130), Expect = 9e-06
Identities = 74/360 (20%), Positives = 137/360 (37%), Gaps = 68/360 (18%)

Query: 1   MKFSKKYIAAGSAVIVSLSLCA--------YALNQHRSQENKDNNRVSYVDGSQSSQKS- 51
            MK K I G + L+L A   Y +   +S++ +N++ + S+ + K+
Sbjct: 1   MKTKKVIILVGLLLLSSQLTLIACQSRGNGTYPIKTKQSRKGMTSNKIKPIKKSKKTNKTH 60

Query: 52   ENLTPDQVSQKEG--IQAEQIVIKITDQGYVTSHGDHYHY----------YNGKVPYDAL 99
            + +   +G + + + ++ TDQG V H H H+   +   +P A
Sbjct: 61   KGVAGVDFPTDDGFILTKDSKILSKTDQGIVVDHDGHSHFIFYADLKGSPFEYLIPKGAS 120

Query: 100 FSEELLMK-----------DPN--YQLKDADIVNEVKGGYIIKVDGKYYVYLKDAAHADN 146
            ++ + +   DP+ Y+   ADIV E GY ++ D ++ LK +
Sbjct: 121 LAKPAVAQQAASQGTSKVADPHHHYEFNPADIVAEDALGYTVRHDDHFHYILKSSLSG-- 178

Query: 147 VRTKDEINRQKQEHVKDNEKVNSNVAVARSQG----RYTTNDGYVFNPADIIEDTGNAYI 202
              +   Q ++   + +S V+ A + G   + T+DG+ FN I+ T + +
Sbjct: 179 -----QTQAQAKQVATRLPQTSSPVSTATANGIPGLHFPTSDGFQFNGQGIVGVTKDNIL 233

Query: 203 VPHGGHYHYIPXXXXXXXXXXXXXXXXXXXXXNMQPSQLSYSSTASDNNTQSVAKGSTSKPA 262
            V H GH H I   P++ +   A ++T +++   +
Sbjct: 234 VDHDGHLHPI-----SFADLRQGGWAHVADQYDPAKKA-EKPAETHHTPELSE----REK 283

Query: 263 NKSENLQSLLKELYDSPSA-QRYSESDGLVFDPAKIISRTPNGVAIPHGDHYHFIPYSKL 321
            E L L ++L PS +R   DG +   G+ PH DH H + S +
Sbjct: 284 EYQEKLAYLAEKLGIDPSTIKRVETQDGKL------------GLEYPHHDHAHVLMLSDI 331

```
tr    Q99Z76      Putative internalin A [inlA] [Streptococcus         792 AA
                  pyogenes]                                           align
```

```
Score = 91.7 bits (226), Expect = 7e-17
Identities = 89/343 (25%), Positives = 137/343 (38%), Gaps = 58/343 (16%)
```

```
Query: 237  SQLSYSSTASDNN------TQSVAKGSTS---KPANKSENLQSLLKELYDSPSAQRYSES 287
            SQL+ +  S N       T+   KG TS   KP  KS+     K +    +    +
Sbjct: 16   SQLTLIACQSRGNGTYPIKTKQSRKGMTSNKIKPIKKSKKTNKTHKGV----AGVDFPTD 71

Query: 288  DGLVFDP-AKIISRTPNGVAIPHGDHYHFIPYSKL--SALEEKIARRVPISGTGSTVSTN 344
            DG +    +KI+S+T  G+ + H  H HFI Y+ L   S   E  I+    +
Sbjct: 72   DGFILTKDSKILSKTDQGIVVDHDGHSHFIFYADLKGSPFEYLIPKGASL---------- 121

Query: 345  AKPNEVVXXXXXXXXXXXXXXXXXXXKELSSASDGYIFNPKDIVEETATAYIVRHGDHFHYIP 404
            AKP                      .+++    Y FNP DIV E A  Y VRH DHFHYI
Sbjct: 122  AKP---------AVAQRAASQGTSKVADPHHHYEFNPADIVAEDALGYTVRHDDHFHYIL 172

Query: 405  KSNQIGQPTLPNNSLAT--PSPSLPIN-------PGISHEKHEEDGYGFDANRIIAEDES 455
            KS+  GQ       +AT  P  S  ++       PG+       DG+ F+   I+    +
Sbjct: 173  KSSLSGQTQAQAKQVATRLPQTSSLVSTATANGIPGLHFP--TSDGFQFNGQGIVGVTKD 230

Query: 456  GFIMSHGNHNHYFFKKDLTEEQIKAAQKHLEEVKTSHNGLDSLSSHEQDYPGNAKEMKDL 515
            ++  H  H H      DL +      + K + +    +H+          E+ +
Sbjct: 231  SILVDHDGHLHPISFADLRQGGWAHVADQYDPAKKAEKPAE---THQ------TPELSER 281

Query: 516  DKKIEEKIAGIMKQYGVKRESI---VVNKEKNAIIYPHGDHHH 555
            +K+ +EK+A + ++  G+    +I       K   + YPH DH H
Sbjct: 282  EKEYQEKLAYLAEKLGIDPSTIKRVETQDGKLGLEYPHHDHAH 324
```

```
Score = 55.1 bits (131), Expect = 7e-06
Identities = 74/360 (20%), Positives = 137/360 (37%), Gaps = 68/360 (18%)
```

```
Query: 1    MKFSKKYIAAGSAVIVSLSLCA--------YALNQHRSQENKDNNRVSYVDGSQSSQKS- 51
            MK  K  I G +   L+L A         Y +   +S++   +N++  + S+ + K+
Sbjct: 1    MKTKKVIILVGLLLSSQLTLIACQSRGNGTYPIKTKQSRKGMTSNKIKPIKKSKKTNKTH 60

Query: 52   ENLTPDQVSQKEG--IQAEQIVIKITDQGYVTSHGDHYHY----------YNGKVPYDAL 99
            + +        +G  +   + + ++ TDQG V  H  H H+          +    +P  A
Sbjct: 61   KGVAGVDFPTDDGFILTKDSKILSKTDQGIVVDHDGHSHFIFYADLKGSPFEYLIPKGAS 120

Query: 100  FSEELLMK----------DPN--YQLKDADIVNEVKGGYIIKVDGKYYVYLKDAAHADN 146
            ++  + +           DP+  Y+   ADIV E   GY ++ D  ++ LK +
Sbjct: 121  LAKPAVAQRAASQGTSKVADPHHHYEFNPADIVAEDALGYTVRHDDHFHYILKSSLSG-- 178

Query: 147  VRTKDEINRQKQEHVKDNEKVNSNVAVARSQG----RYTTNDGYVFNPADIIEDTGNAYI 202
               +   Q ++       + +S V+ A + G       + T+DG+ FN  I+  T ++ +
Sbjct: 179  -----QTQAQAKQVATRLPQTSSLVSTATANGIPGLHFPTSDGFQFNGQGIVGVTKDSIL 233

Query: 203  VPHGGHYHYIPXXXXXXXXXXXXXXXXXXXXXNMQPSQLSYSSTASDNNTQSVAKGSTSKPA 262
            V H GH H I                     P++ +   A  + T +++        +
Sbjct: 234  VDHDGHLHPI-----SFADLRQGGWAHVADQYDPAKKA-EKPAETHQTPELSE----REK 283
```

```
Query:  263  NKSENLQSLLKELYDSPSA-QRYSESDGLVFDPAKIISRTPNGVAIPHGDHYHFIPYSKL 321
                  E L  ++L    PS  +R    DG +             G+  PH DH H +  S +
Sbjct:  284  EYQEKLAYLAEKLGIDPSTIKRVETQDGKL------------GLEYPHHDHAHVLMLSDI 331
```

tr Q6HCJ0 **Cell division protein [ftsK] [Bacillus thuringiensis (subsp.** 1338 AA
           konkukian)]

                                                                          align

 Score = 46.6 bits (109), Expect = 0.002
 Identities = 52/223 (23%), Positives = 96/223 (42%), Gaps = 23/223 (10%)

```
Query:  789  AYLDNQSTYIVEVPILEKENQTDKPSILPQFKRNKAQENSKLDEKVEEPKTSEKVEKEKL 848
             A +  + + + E P++E++    ++ SI+ +  +  A+E S ++E+    +TS   E+
Sbjct:  585  APIAEEQSVVEEAPVVEEQPVAEETSIVEE--QPAAEETSVVEEQPVAEETSVVEEQPVA 642

Query:  849  SETGNSTSNSTLEEVPTVDP---VQEKVAKFAESYGMKLENVLFNMDGT--IELYLPSGE 903
               ET       EE P V+    VQ++  K +  +    V+   D   +E +
Sbjct:  643  EETSIVEEQPVAEEAPVVEEQPVVQKEEPKREKKRHVPFNVVMLKQDRARLMERHASRTN 702

Query:  904  VIKKNMADFTGEAPQGNGENKPSENGKVSTGTVENQPTENKPADSLPEAP--NEKPVKPE 961
             ++ +M++     P   E KP  K    VE Q  E KP   +    P    EKPV+
Sbjct:  703  AMQPSMSERVENKPVHQVEEKPQVEEKPVQQVVEPQ-VEEKPMQQVVVEPQVEEKPVQQ- 760

Query:  962  NSTDNGMLNPEGNVGSDPM----LDPALEEAPAVDPVQEKLEK 1000
                   ++ P+  V   PM     ++P +EE P    V+ ++E+
Sbjct:  761  ------VVEPQ--VEEKPMQQVVVEPQVEEKPVQQVVEPQVEE 795
```

sp P45386   **Immunoglobulin A1 protease precursor (EC 3.4.21.72) (IGA1** 1849
   IGA4_HAEIN **protease)**   AA
              [iga] [Haemophilus influenzae]   align

 Score = 46.2 bits (108), Expect = 0.003
 Identities = 54/226 (23%), Positives = 88/226 (38%), Gaps = 30/226 (13%)

```
Query:  788  NAYLDNQSTYIVEVPILEKENQTDKPSILPQFKRNKAQENSKLDEKVEEPKTSEKVEKEK 847
             +A +N+    VE P+      T+      Q +    A+        +EE   T+   E
Sbjct: 1022  SAQSNNEEIARVETPVPPPAPATESAIASEQPETRPAET---AQPAMEETNTANSTETAP 1078

Query:  848  LSETGNSTSNSTLEEVPTVDPVQEKVAKFAESYGMKLENVLFNMDGTIELYLPSGEVIKK 907
              S+T   T N   E VP+      E   K AE+   + E V N      E    +GEV K+
Sbjct: 1079  KSDTATQTENPNSESVPS-----ETTEKVAENPPQENETVAKNEQEATEPTPQNGEVAKE 1133

Query:  908  NMADF-----TGEAPQGNGENKPSENGKVSTGTVENQPT------ENKPADSLPEAPNEK 956
             +        T EA Q  G+ + ++      T   +++PT        EN+P  ++ ++  +K
Sbjct: 1134  DQPTVEANTQTNEATQSEGKTEETQ-----TAETKSEPTESVTVSENQPEKTVSQSTEDK 1188

Query:  957  PV-----KPENSTDNGMLNPEGNVGSDP-MLDPALEEAPAVDPVQE 996
              V     K + T+    P+      P    +PA EE P    +E
Sbjct: 1189  VVVEKEEKAKVETEETQKAPQVTSKEPPKQAEPAPEEVPTDTNAEE 1234
```

```
tr  Q8MNS0  Hypothetical protein F12F3.2 [F12F3.2] [Caenorhabditis      6328
            elegans]                                                    AA
                                                                        align

  Score = 46.2 bits (108), Expect = 0.003
  Identities = 53/260 (20%), Positives = 101/260 (38%), Gaps = 34/260 (13%)

Query:  791  LDNQSTYIVEVPILEKENQTDKPSILPQFKRNKAQENSKLDEKVEEPKTSEKVEKEKLSE  850
             L+ Q     E+ ++E + + +K         AQE S+L+++ ++    +K+E E  S+
Sbjct: 1368  LEEQKKKAAEIALIEIQKEQEK----------LAQEQSRLEDEAKKSAEKQKLESETKSK  1417

Query:  851  TGNSTSNSTLEEVPTVDPVQEKVAKFAESYGMKLENVLFNMDGTIELYLPSGEVIKKNM-  909
                 +++E P   +++K  K  S   K ++    +D    L     V KK +
Sbjct: 1418  QTEEAPKESVDEKPKKKVLKKKTEKSDSSISQKSKSAKSTVDAAETLESDFNLVEKKTVQ  1477

Query:  910  -----------ADFTGEAPQGNGENKPSENGKVSTGTVENQPTENKPADSLPEAPNEKPV  958
                        A   +  Q  E    ++G    T + +P   KP DS   EA  +K V
Sbjct: 1478  KVEQSPDESTSATIKRDPAQKTEEISKQDDGDEKKTTTDGKPP--KPEDS--EATPKKRV  1533

Query:  959  KPENSTDNGMLNPEGNVGSDPMLDPALEEAPAVDPVQEKLEKFTASYGLGLDSVIFNMDG  1018
             + + + + + + ++    L  +EE P   +++K EK        DSVI
Sbjct: 1534  VKKKTQKSDSVASDASLADVSKLSDDVEEKPKKKVLKKKTEK--------SDSVISETSS  1585

Query: 1019  TIELRLPSGEVIKKNLSDLI  1038
             ++   S E+  +    +I
Sbjct: 1586  VDTIKPESVEIPTEKAEQMI  1605
```

```
tr   Q8ISF7      2MDa_1 protein [isof] [Caenorhabditis elegans]     18534 AA
                                                                     align

  Score = 46.2 bits (108), Expect = 0.003
  Identities = 53/260 (20%), Positives = 101/260 (38%), Gaps = 34/260 (13%)

Query:   791  LDNQSTYIVEVPILEKENQTDKPSILPQFKRNKAQENSKLDEKVEEPKTSEKVEKEKLSE  850
              L+ Q     E+ ++E + + +K         AQE S+L+++ ++    +K+E E  S+
Sbjct: 13574  LEEQKKKAAEIALIEIQKEQEK----------LAQEQSRLEDEAKKSAEKQKLESETKSK  13623

Query:   851  TGNSTSNSTLEEVPTVDPVQEKVAKFAESYGMKLENVLFNMDGTIELYLPSGEVIKKNM-  909
                  +++E P   +++K  K  S   K ++    +D    L     V KK +
Sbjct: 13624  QTEEAPKESVDEKPKKKVLKKKTEKSDSSISQKSKSAKSTVDAAETLESDFNLVEKKTVQ  13683

Query:   910  -----------ADFTGEAPQGNGENKPSENGKVSTGTVENQPTENKPADSLPEAPNEKPV  958
                         A   +  Q  E    ++G    T + +P   KP DS   EA  +K V
Sbjct: 13684  KVEQSPDESTSATIKRDPAQKTEEISKQDDGDEKKTTTDGKPP--KPEDS--EATPKKRV  13739

Query:   959  KPENSTDNGMLNPEGNVGSDPMLDPALEEAPAVDPVQEKLEKFTASYGLGLDSVIFNMDG  1018
              + + + + + + ++    L  +EE P   +++K EK        DSVI
Sbjct: 13740  VKKKTQKSDSVASDASLADVSKLSDDVEEKPKKKVLKKKTEK-------SDSVISETSS   13791

Query:  1019  TIELRLPSGEVIKKNLSDLI  1038
              ++   S E+  +    +I
Sbjct: 13792  VDTIKPESVEIPTEKAEQMI  13811
```

```
tr    Q8ISF6      2MDa_2 protein [isof] [Caenorhabditis elegans]      18519 AA
                                                                      align

  Score = 46.2 bits (108), Expect = 0.003
  Identities = 53/260 (20%), Positives = 101/260 (38%), Gaps = 34/260 (13%)

Query: 791   LDNQSTYIVEVPILEKENQTDKPSILPQFKRNKAQENSKLDEKVEEPKTSEKVEKEKLSE 850
             L+ Q      E+ ++E + + +K            AQE S+L+++ ++    +K+E E  S+
Sbjct: 13574 LEEQKKKAAEIALIEIQKEQEK---------LAQEQSRLEDEAKKSAEKQKLESETKSK 13623

Query: 851   TGNSTSNSTLEEVPTVDPVQEKVAKFAESYGMKLENVLFNMDGTIELYLPSGEVIKKNM- 909
                    +++E P    +++K  K   S   K ++     +D     L     V KK +
Sbjct: 13624 QTEEAPKESVDEKPKKKVLKKKTEKSDSSISQKSKSAKSTVDAAETLESDFNLVEKKTVQ 13683

Query: 910   ----------ADFTGEAPQGNGENKPSENGKVSTGTVENQPTENKPADSLPEAPNEKPV 958
                       A   +  Q   E   ++G    T + +P    KP DS   EA  +K V
Sbjct: 13684 KVEQSPDESTSATIKRDPAQKTEEISKQDDGDEKKTTTDGKPP--KPEDS--EATPKKRV 13739

Query: 959   KPENSTDNGMLNPEGNVGSDPMLDPALEEAPAVDPVQEKLEKFTASYGLGLDSVIFNMDG 1018
             + + + + + ++         L  +EE P   +++K EK          DSVI
Sbjct: 13740 VKKKTQKSDSVASDASLADVSKLSDDVEEKPKKKVLKKKTEK-------SDSVISETSS 13791

Query: 1019  TIELRLPSGEVIKKNLSDLI 1038
             ++  S E+ +    +I
Sbjct: 13792 VDTIKPESVEIPTEKAEQMI 13811


sp    P16053      Neurofilament triplet M protein (160 kDa neurofilament     857
      NFM_CHICK    protein)                                                 AA
                  (Neurofilament medium polypeptide) (NF-M) [NEFM] [Gallus    align
                  gallus (Chicken)]

  Score = 44.7 bits (104), Expect = 0.009
  Identities = 58/236 (24%), Positives = 85/236 (35%), Gaps = 39/236 (16%)

Query: 760   GEIKLPIPKLNQGTTRTAGNKIPVTFMANAYLDNQSTYIVEVPIL-----EKENQTDKPS 814
             G I PI   Q+   A KI T+    L Q ++ E+     EK    D  S
Sbjct: 411   GSITGPIFTHRQPSVTIASTKIQKTKIEPPKLKVQHKFVEEIIEETKVEDEKSEMEDALS 470

Query: 815   ILPQFKRNKAQENSKLDEKVEEPKTSEKVEKEKLSETGNSTSNSTLEEVPTVDPVQEKVA 874
              + +     KAQE + +EK EE   E+   EK +E           EE        +E+ A
Sbjct: 471   AIAEEMAAKAQEEEQEEEKAEEEAVEEEAVSEKAAEQAAEEEEKEEEEAE-----EEEAA 525

Query: 875   K--FAESYGMKLENVLFNMDGTIELYLPSGEVIKKNMADFTGEAPQGNGENKPSENGKVS 932
             K   AE G K E +   +G             EA +  E K   GK
Sbjct: 526   KSDAAEEGGSKKEEIEEKEEGE------------------EAEEEEAEAK----GKAE 561

Query: 933   TGTVENQPTENKPADSLPEAPNEKPVKPENSTDNGMLNPEGNVGSDPMLDPALEEA 988
             + +  ++ PA S P++P + PV +       +      VG D  + A E+A
Sbjct: 562   EAGAKVEKVKSPPAKSPPKSPPKSPVTEQ---AKAVQKAAAEVGKDQKAEKAAEKA 614


tr    Q97QP7  Immunoglobulin A1 protease [SP1154] [Streptococcus        2004
              pneumoniae]                                               AA
                                                                        align
```

Score = 44.3 bits (103), Expect = 0.012
Identities = 61/251 (24%), Positives = 99/251 (39%), Gaps = 46/251 (18%)

```
Query:  800  EVPILEKENQTDKPSILPQFKRNKAQENSKLDEKVEEPKTSEKVEKEKLSETGNSTSNST 859
             E P+   E T+ SI    +  ES ++  P TS K  E  S  +ST++
Sbjct:  512  ETPVNPNEGTTEGTSIQEAENPVQPAEESTTNSEKVSPDTSSKNTGEVSSNPSDSTTSVG 571

Query:  860  LEEVPTVDPVQEKVAKFAESYGMKLENVLFN-MDGTIELYLPSGEVIKKNMADFTGEAPQ 918
                 P + +      E+  +E V N +GT+E          ++   E P
Sbjct:  572  ESNKPEHNDSKN------ENSEKTVEEVPVNPNEGTVE----------GTSNQETEKPV 614

Query: ·919  GNGENKPSENGKV---STGTVENQPTENKP--ADSLPEAPNEKPVKPENS----TDNGML 969
             E  + +GK+    +TG V N+P+++KP    +S    N    KPENS    ++NG
Sbjct:  615  QPAEETQTNSGKIANENTGEVSNKPSDSKPPVEESNQPEKNGTATKPENSGNTTSENGQT 674

Query:  970  NPE---GNVGSDPMLDPALEEAPAVDPVQEKLEKFTASYGLGLDSVIFNMDGTIELRLPS 1026
             PE   GN  D  +   +   + ++++ E       L  D  +   T+ELR
Sbjct:  675  EPEPSNGNSTEDVSTESNTSNSNGNEEIKQENE-------LDPDKKVEEPEKTLELR--- 724

Query: 1027  GEVIKKNLSDL 1037
                     N+SDL
Sbjct:  725  ------NVSDL 729
```

tr <u>Q8IB63</u> **Hypothetical protein PF08_0035 [PF08_0035] [Plasmodium**   1176
             falciparum   AA
             (isolate 3D7)]   <u>align</u>

Score = 44.3 bits (103), Expect = 0.012
Identities = 56/237 (23%), Positives = 94/237 (39%), Gaps = 27/237 (11%)

```
Query:  778  GNKIPVTFMANAYLDNQSTYIVEVPILEKENQTDKPSILPQFKRNKAQENSKLDEKVEEP 837
             GNK     N  D  T  E  +E E  +K +I+ Q     Q+N KLDE++ E
Sbjct:  205  GNKKNKNKNKNKNSDKSETNDEEEVTVEME---EKENIIEQLNE---QQNEKLDEQLNE- 257

Query:  838  KTSEKVEKEKLSETGNSTSNSTLEEVPTVDPVQEKVAKFAESYGMKLENVLFNMDGTIEL 897
             + +E + E+ +E +   N   +E P + ++     F E    N   F+
Sbjct:  258  QPNENFD-EQPNENFDEQPNENFDEQPNENFDEQ-----PNENFDEQPNENF 311

Query:  898  YLPSGEVIKKNMADFTGEAPQGNGENKPSEN-GKVSTGTVENQPTEN---KPADSLPEAP 953
                 E  +  +  E P  N + +P+EN  +    + QP EN  +P ++  E P
Sbjct:  312  DEQPNENFDEQPNENFDEQPNENFDEQPNENFDEQPNENFDEQPNENFDEQPNENFDEQP 371

Query:  954  NEKPVKPENSTDNGMLNPEGNVGSDPMLDPA---------LEEAPAVDPVQEKLEK 1000
             NE  + N+T+   N   V ++ ++D          +EE    D + E+ EK
Sbjct:  372  NENSDEQSNNTNKDEPNTSNYVNTENVIDQGKGNNSECGNIVEEQNNGDSLDERKEK 428
```

Score = 40.0 bits (92), Expect = 0.23
Identities = 43/193 (22%), Positives = 76/193 (39%), Gaps = 9/193 (4%)

```
Query:  805  EKENQTDKPSILPQFKRNKAQENSKLDEKVEEPKTSEKVEKEKLSETGNSTSNSTLEEVP 864
             +K+ +  K      +NK ++ K +  EE  T E  EKE + E N   N   L+E
Sbjct:  196  QKKKKKKKKGNKKNKNKNKNKNSDKSETNDEEEVTVEMEEKENIIEQLNEQQNEKLDEQL 255

Query:  865  TVDPVQEKVAKFAESYGMKLENVLFNMDGTIELYLPSGEVIKKNMADFTGEAPQGNGENK 924
```

```
                P +      +  E++    +  N F+                E  +   +   E P N + +
Sbjct: 256 NEQPNENFDEQPNENFD-EQPNENFDEQPNENFDEQPNENFDEQPNENFDEQPNENFDEQ 314

Query: 925 PSEN-GKVSTGTVENQPTEN---KPADSLPEAPN----EKPVKPENSTDNGMLNPEGNVG 976
                P+EN +       + QP EN  +P ++  E PN    E+P +   +   N   + + N
Sbjct: 315 PNENFDEQPNENFDEQPNENFDEQPNENFDEQPNENFDEQPNENFDEQPNENFDEQPNEN 374

Query: 977 SDPMLDPALEEAP 989
                SD   +   ++ P
Sbjct: 375 SDEQSNNTNKDEP 387
```

tr <u>Q869E1</u> **Similar to Xenopus laevis (African clawed frog). DNA ligase I** 1192
             (EC   AA
             6.5.1.1) (Polydeoxyribonucleotide synthase   <u>align</u>
             [Dictyostelium discoideum (Slime mold)]

 Score = 44.3 bits (103), Expect = 0.012
 Identities = 46/176 (26%), Positives = 76/176 (43%), Gaps = 12/176 (6%)

```
Query: 820 KRNKAQENSKLD---EKVEEPKTSEKVEKEKLSETGNSTSNSTL-----EEVPTVDPVQE 871
                K + AQE +KL   +K  E    EKVEK++  ET    +S       EE  +   E
Sbjct: 23  KISNAQELNKLKTSPKKKREAVVKEKVEKKEKKETKPKRKSSKKNKEEEEEEEQEEQDGE 82

Query: 872 KVAKFAESYGMKLENVLFNMDGTIELYLPSGEVIKKNMADFTGEAPQGNGENKPSENGKV 931
                + +   E Y  + E+  +++G  E+ L   E  K       + +      ENK S++
Sbjct: 83  EEQEEEEEYQQQDEEIEEDINGEEEMELDENEKEKNKKKKQSLKTK----ENKESKSSSS 138

Query: 932 STGTVENQPTENKPADSLPEAPNEKPVKPENSTDNGMLNPEGNVGSDPMLDPALEE 987
                S  T+EN+ T+     S  ++ N K +K +    D+     + N   D  LD  L++
Sbjct: 139 SKKTIENKETKKPEKQSSKQSNNLKRLKRKKMDDDEEDEEDENKTDDNDLDDMLDD 194
```

tr <u>Q9VC00</u> **CG13648-PA [CG13648] [Drosophila melanogaster (Fruit fly)]** 2768 AA

                                                                    <u>align</u>

 Score = 43.9 bits (102), Expect = 0.016
 Identities = 57/218 (26%), Positives = 89/218 (40%), Gaps = 29/218 (13%)

```
Query: 773  TTRTAGNKIPVTFMANAYLDNQSTYIVEVPI--LEKENQTDKPSILPQFKRNKAQENSKL 830
                 T   AG KIP + +         ++ V P  +EKE  T+ P+  P     ++ ENS
Sbjct: 1651 TVAPAGEKIPTSSITPDEEPTATSAPVAKPDEDVEKETSTEIPTDAPA--SSEEDENSST 1708

Query: 831  DEKVEE-----PKTSEKVEKEKLSETGNSTSNSTLEEVPTVDPVQEKVAKFAESYGMKLE 885
                 D+   E     P+T + +E    G + + +T +EVP V + E+V   AE      E
Sbjct: 1709 DQIPSEVPEKKPETPAQTPEEG-DIVGATAAPTTSDEVPPVQRLPEEV--LAEIPQPSTE 1765

Query: 886  NVLFNMDGTIELYLPSGEVIKKNMADFTGEA----PQGNGENKPSENGKVSTGTVENQPT 941
                 +   D T    PS + +  + + EA    P    + KP+E        E +P
Sbjct: 1766 TGIKQQDETTAA--PSIDRKEPYVTEIDEEATTVAPISEKDEKPTE--------EEKPV 1814

Query: 942  ENKPADSLP--EAPNEKPVKPENSTDNGMLNPEGNVGS 977
                 E KP    P  E     EKP++ +  ST+  +       GS
Sbjct: 1815 EQKPTGEEPSEEEEKEKPIEQDVSTEGPVSTEASEAGS 1852
```

tr <u>Q839C3</u> **N-acetylmuramoyl-L-alanine amidase, family 4 [EF0252]**   503
              [Enterococcus   AA
              faecalis (Streptococcus faecalis)]   <u>align</u>

  Score = 43.5 bits (101), Expect = 0.021
  Identities = 41/184 (22%), Positives = 68/184 (36%), Gaps = 9/184 (4%)

Query: 822 NKAQENSKLDEKVEEPKTSEKVEKEKLSETGNSTSNSTLEEVPTVDPVQEKVAKFAESYG 881
             N ENS + E   E T+E E   SET +T ST E + + +   + E+ G
Sbjct: 30   NLVAENSSVAETTAEATTAEATETTATSETTEATEESTTETESSTESSESATTESTETSG 89

Query: 882 MKLENVLFNMDGTIELYLPSGEVIKKNMADFTGEAPQGNGENKPSENGKVSTGTVENQPT 941
             + +   +   T   + E   ++ T E+ +   +E+   T T   +
Sbjct: 90   TETTDSTTDSTST-----STTESTTDSTSTSTTESTTDSTSTSTTESSTTPTTTPSSSKE 144

Query: 942 ENKPADSLPEAPN-EKPVKPENSTDNGMLNPEGNVGSDPMLDPALEEAPAVDPVQEKLEK 1000
             + KP S E+   KPV P   + + P +   ++ P E   + E E
Sbjct: 145 QPKPGTSTSESKQPAKPVTPTAPAEKPVEQPAASTPQPEIVPPVTNETVG---LVEDDET 201

Query: 1001 FTAS 1004
             FT S
Sbjct: 202 FTVS 205

tr   <u>O73793</u>   **Neurofilament medium subunit [Serinus canaria**   487 AA
                    (Canary)]   <u>align</u>

  Score = 43.1 bits (100), Expect = 0.027
  Identities = 57/236 (24%), Positives = 85/236 (35%), Gaps = 39/236 (16%)

Query: 760 GEIKLPIPKLNQGTTRTAGNKIPVTFMANAYLDNQSTYIVEVPIL-----EKENQTDKPS 814
             G I PI   Q +   A KI T +   L Q ++ E+   EK   D S
Sbjct: 41   GSITGPIFTHRQPSVTIASTKIQKTKIEPPKLKVQHKFVEEIIEETKVEDEKSEMEDALS 100

Query: 815 ILPQFKRNKAQENSKLDEKVEEPKTSEKVEKEKLSETGNSTSNSTLEEVPTVDPVQEKVA 874
             + +   KAQE + +EK EE   E+   EK +E   EE   +E+ A
Sbjct: 101 AIAEEMAAKAQEEEQEEEKAEEEAVEEEAVSEKAAEQAAEEEEKEEEEAE-----EEEAA 155

Query: 875 K--FAESYGMKLENVLFNMDGTIELYLPSGEVIKKNMADFTGEAPQGNGENKPSENGKVS 932
             K AE G + E +   +G   EA +   E K   GK
Sbjct: 156 KSDAAEEGGSEKEEIEEKEEGE-------------------EAEEEEAEAK----GKAE 191

Query: 933 TGTVENQPTENKPADSLPEAPNEKPVKPENSTDNGMLNPEGNVGSDPMLDPALEEA 988
             + + ++ PA S P++P + PV +   +   VG D   + A E+A
Sbjct: 192 EAGAKVEKVKSPPAKSPPKSPPKSPVTEQ---AKAVQKAAAEVGKDQKAEKAAEKA 244

tr <u>Q9FN97</u> **Transposon protein-like [Arabidopsis thaliana (Mouse-ear**   1089
              cress)]   AA
                                                                           <u>align</u>

```
 Score = 43.1 bits (100), Expect = 0.027
 Identities = 31/131 (23%), Positives = 56/131 (42%), Gaps = 2/131 (1%)

Query:  19  SLCAYALNQHRSQENKDNNRVSYVDGSQSSQKSENLTPDQVSQKEGIQAEQIVIKITDQG 78
            + C Y  ++H  +N+ +R  Y  G+Q        +  PD SQ       ++  I + D+
Sbjct: 792  TFCTYYFDRHIKTKNRAGDR-HYDGGNQEDTHEFDGVPDIFSQAGRDSGKESEIWLQDKD 850

Query:  79  YVTSHGDHYHYYNGKVPYDALFSEELLMKDPNYQLKDADIVNEVK-GGYIIKVDGKYYVY 137
            Y +H        +   P++ LF E L+ +P   KD + + E +  ++ K D  Y +
Sbjct: 851  YHIAHRYILRNCDQLRPFERLFDESLIAANPGISEKDLNELREKQYSSWLKKYDNSYPEW 910

Query: 138  LKDAAHADNVR 148
            L   H   V+
Sbjct: 911  LLSIVHGPMVK 921
```

tr <u>Q963T1</u> **Glutamate-rich protein (Fragment) [GLURP] [Plasmodium**   1078
                reichenowi]   AA
                                                                                        <u>align</u>

```
 Score = 43.1 bits (100), Expect = 0.027
 Identities = 48/182 (26%), Positives = 82/182 (44%), Gaps = 21/182 (11%)

Query: 791  LDNQSTYIVE----VPILEKENQTDKPSILPQFKRN-KAQENSKLDEKVEEPKTSEKVEK 845
            LDN+  +I E      V +L+ EN  +   + PQ K N ++  E   +D ++  P+  EK E
Sbjct: 360  LDNEKEHIDESQHNVNVLQ-ENNINNHQLEPQEKANIESFEPKNIDSEIILPENVEKEEI 418

Query: 846  EKLSETGNSTSNSTLEEVPTVDPVQEKVAKFAESYGMKLENVLFNMDGTIELYLPSGEVI 905
             +     ++ TLE+  +    +E V++             N   T+E       +
Sbjct: 419  VVDVPSPKHFNHETLEQETSESEHEEAVSEK------------NAHETVEHEEAVSQES 465

Query: 906  KKNMADFTGEAPQGNGENKPSENGKVSTGTVENQPTENKPADSLPEAPNEKPVKPENSTD 965
               AD G   Q N  N+ +EN  V +    E++P  EN+ +  SL EA  E+ V  +N+ +
Sbjct: 466  NPEKADNDGNVSQ-NSNNELNENEFVESEKSEHEPAENEES-SLEEAHQEEIVPEQNNQE 523

Query: 966  NG 967
            +G
Sbjct: 524  SG 525
```

tr   <u>Q07594</u>   **K2 protein (Fragment) [Entamoeba histolytica]** 242 AA
                                                                        <u>align</u>

```
 Score = 43.1 bits (100), Expect = 0.027
 Identities = 50/203 (24%), Positives = 80/203 (38%), Gaps = 32/203 (15%)

Query: 803  ILEKENQTDKPSILPQFKRNKAQENSKLDEKVEEPKTSEKVEKEKLSETGNSTSNSTLEE 862
            IL+ + +     I  F +N+A     EK+EE K  E+ EK  S TG+   N++  E
Sbjct:  14  ILDIDQEVKDTKIYSVFLKNEASP-----EKLEENKEDEEKEKSSSSNTGDEPDNNS--E 66

Query: 863  VPTVDPVQEKVAKFAESYGMKLENVLFNMDGTIELYLPSGEVIKKNMADFTGEAPQGNGE 922
            + D  E  A  +   K+E              S +  + + +D    P+ +
Sbjct:  67  AKSDD---EPEASSNDKSDDKVE------------ASSSDKPEASSSDKPDNKPEASSS 110

Query: 923  NKPSENGKVSTGTVENQPTENKPADSLPEAPNEKPV-----KPE-NSTDNGMLNPEGNVG 976
```

```
             +KP   +              +   +NKP +S  + P+ KP      KPE +S+D     PE +
Sbjct: 111 DKPEASSSDKPEASSSDKPDNKPEESSSDKPDNKPEASSSDKPEASSSDKPDNKPEASSS 170


Query: 977 SDPMLDPALEEAPAVDPVQEKLE 999
              P     P    EA + D    K E
Sbjct: 171 DKPDNKP---EASSSDKPDNKPE 190
```

```
tr    Q6PK21          OGFR protein [Homo sapiens (Human)]          513 AA
                                                                   align
```

```
 Score = 43.1 bits (100), Expect = 0.027
 Identities = 26/100 (26%), Positives = 46/100 (46%), Gaps = 8/100 (8%)

Query: 900 PSGEVIKKNMADFTGEAPQGNGENKPSENGKVSTGTVENQPTENKPADSLPEAPNEKPVK 959
           P+G+   ++ ++  G  P G   ++P+E+   + G     PT ++PA+S  E P  +P
Sbjct: 385 PAGDEPAESPSETPGPRPAGPAGDEPAESPSETPGPSPAGPTRDEPAESPSETPGPRPAG 444

Query: 960 P------ENSTDNGMLNPEGNVGSDPMLDPALEEAPAVDP 993
           P       E+ ++      P G  G +P   P+   E P   P
Sbjct: 445 PAGDEPAESPSETPGPRPAGPAGDEPAESPS--ETPGPSP 482
```

```
 Score = 35.4 bits (80), Expect = 5.7
 Identities = 21/82 (25%), Positives = 35/82 (42%), Gaps = 6/82 (7%)

Query: 914 GEAPQGNGENKPSENGKVSTGTVENQPTENKPADSLPEAPNEKPVKP------ENSTDNG 967
           G +P G   ++P+E+   + G      P ++PA+S  E P   P        E+ ++
Sbjct: 379 GPSPAGPAGDEPAESPSETPGPRPAGPAGDEPAESPSETPGPSPAGPTRDEPAESPSETP 438

Query: 968 MLNPEGNVGSDPMLDPALEEAP 989
             P G  G +P   P+     P
Sbjct: 439 GPRPAGPAGDEPAESPSETPGP 460
```

```
sp_vs  Q9NZT2-2  Splice isoform 2 of Q9NZT2 [OGFR] [Homo sapiens (Human)]  657 AA

                                                                         align
```

```
 Score = 43.1 bits (100), Expect = 0.027
 Identities = 26/100 (26%), Positives = 46/100 (46%), Gaps = 8/100 (8%)

Query: 900 PSGEVIKKNMADFTGEAPQGNGENKPSENGKVSTGTVENQPTENKPADSLPEAPNEKPVK 959
           P+G+   ++ ++  G  P G   ++P+E+   + G     PT ++PA+S  E P  +P
Sbjct: 529 PAGDEPAESPSETPGPRPAGPAGDEPAESPSETPGPSPAGPTRDEPAESPSETPGPRPAG 588

Query: 960 P------ENSTDNGMLNPEGNVGSDPMLDPALEEAPAVDP 993
           P       E+ ++      P G  G +P   P+   E P   P
Sbjct: 589 PAGDEPAESPSETPGPRPAGPAGDEPAESPS--ETPGPSP 626
```

```
 Score = 35.4 bits (80), Expect = 5.7
 Identities = 21/82 (25%), Positives = 35/82 (42%), Gaps = 6/82 (7%)

Query: 914 GEAPQGNGENKPSENGKVSTGTVENQPTENKPADSLPEAPNEKPVKP------ENSTDNG 967
```

```
               G  +P G   ++P+E+    + G     P  ++PA+S  E P    P  P       E+ ++
Sbjct: 523 GPSPAGPAGDEPAESPSETPGPRPAGPAGDEPAESPSETPGPSPAGPTRDEPAESPSETP 582


Query: 968 MLNPEGNVGSDPMLDPALEEAP 989
               P G   G +P   P+     P
Sbjct: 583 GPRPAGPAGDEPAESPSETPGP 604
```

sp <u>Q9NZT2</u>   **Opioid growth factor receptor (OGFr) (Zeta-type opioid**   677
    OGFR_HUMAN **receptor)**   AA
                  (7-60 protein) [OGFR] [Homo sapiens (Human)]   <u>align</u>

 Score = 42.7 bits (99), Expect = 0.036
 Identities = 26/100 (26%), Positives = 45/100 (45%), Gaps = 8/100 (8%)

```
Query: 900 PSGEVIKKNMADFTGEAPQGNGENKPSENGKVSTGTVENQPTENKPADSLPEAPNEKPVK 959
               P+G+   ++ ++   G  P G   ++P+E    + G       PT ++PA+S  E P  +P
Sbjct: 549 PAGDEPAESPSETPGPRPAGPAGDEPAETPSETPGPSPAGPTRDEPAESPSETPGPRPAG 608


Query: 960 P------ENSTDNGMLNPEGNVGSDPMLDPALEEAPAVDP 993
               P      E+ ++     P G   G +P   P+    E P    P
Sbjct: 609 PAGDEPAESPSETPGPRPAGPAGDEPAESPS--ETPGPSP 646
```

 Score = 37.4 bits (85), Expect = 1.5
 Identities = 22/96 (22%), Positives = 42/96 (42%), Gaps = 6/96 (6%)

```
Query: 900 PSGEVIKKNMADFTGEAPQGNGENKPSENGKVSTGTVENQPTENKPADSLPEAPNEKPVK 959
               P+G+   ++ ++   G  P G   ++P+E+    + G       P  ++PA++  E P    P
Sbjct: 529 PAGDEPAESPSETPGPRPAGPAGDEPAESPSETPGPRPAGPAGDEPAETPSETPGPSPAG 588


Query: 960 P------ENSTDNGMLNPEGNVGSDPMLDPALEEAP 989
               P      E+ ++     P G   G +P   P+     P
Sbjct: 589 PTRDEPAESPSETPGPRPAGPAGDEPAESPSETPGP 624
```

tr <u>Q6HBX5</u> **Hypothetical protein [BT9727_4640] [Bacillus thuringiensis**   3471
            (subsp.   AA
            konkukian)]   <u>align</u>

 Score = 42.7 bits (99), Expect = 0.036
 Identities = 52/208 (25%), Positives = 84/208 (40%), Gaps = 28/208 (13%)

```
Query: 808  NQTDKP--SILPQFKRNKAQENSKLDEKVEEPKTSEKVEKEKLSETGNSTSNSTLEEVPT 865
                NQ ++P  SI PQ+K    +  ++  K  EPK  E+VE    + N+  N  L EV
Sbjct: 3132 NQPEEPVVSITPQYKDGMLKAEKEVSNK--EPKLGEEVEYRISFK--NTVENGKLAEVKI 3187


Query: 866  VDPVQEKVAKFAESYGMK--------------LENVLFNMDGTIELYLPSGEVIKKNMA 910
                 D + +   +S  K              EN++  +  I    +V +K
Sbjct: 3188 EDQLPDGLEYVKDSVKAKGAIEVKVENGKLTAKYENIIDTKERNITFKV---KVKEKAGE 3244


Query: 911  DFTGEAPQGNGENKPSE-NGKVSTGTVENQPTENKPADSLPEAPNEKPVKPENSTDNGML 969
                +    A  +G N+P E    +    E +P + K  + PE P E  VKPE+   +
Sbjct: 3245 EIVNRAIVDDGINQPLEPTVSIKPKEPEVKPEDPKEPEVKPEDPKEPEVKPEDPKE-PEV 3303
```

```
Query: 970    NPEGNVGSDPMLDPALEEAPAVDPVQEK 997
              PE     +P  +  P    + P V P   K
Sbjct: 3304   KPED--PKEPEVKPEDPKEPEVKPEDPK 3329
```

tr <u>Q7RQS8</u> **Retinitis pigmentosa GTPase regulator-like protein [PY01014]** 674 AA
         [Plasmodium yoelii yoelii]

<u>align</u>

 Score = 42.7 bits (99), Expect = 0.036
 Identities = 45/193 (23%), Positives = 82/193 (42%), Gaps = 16/193 (8%)

```
Query: 771    QGTTRTAGNKIPVTFMANAYL-DNQSTYIVEVPILEKENQTDKPSILPQFKRNKAQENSK 829
              +   T  +   +P T  N+   +N+S + + P  +    ++P I+ + + +     N
Sbjct: 431    ENPTSLSNTNVPRTGEGNSIQQENRSPNVADRPPTKTSGVENRPMII-ELEVSSDASNFI 489

Query: 830    LDEKVEEPKTSEKVEKEKLSETGNSTSNSTLEEVPTVDPVQEKVAKFAESYGMKLENVLF 889
              LD    E +  E+ E+E+  E +       LE++  +  +E +    E+  ++ E+
Sbjct: 490    LDNDETEEEHDEEPEEEEAEEETDENLQDILEDIEEGED-EEDMEDDDENETLEEES--- 545

Query: 890    NMDGTIELYLPSGEVIKKNMADFTGEAPQGNGENKPSENGKVSTGTVENQPTENKPADSL 949
                 +    SG  + N+ D  EA   GE    EN  ++ T EN  TE+   A
Sbjct: 546    ---NEVTRETSSGGDERPNVTD---EASPTEGETTAQENNSDNSRTNENSTTESTNA--- 596

Query: 950    PEAPNEKPVKPEN 962
              E  N+KP KP+N
Sbjct: 597    -EQKNKKPPKPQN 608
```

tr <u>O77320</u> **Hypothetical protein MAL3P3.3 [MAL3P3.3] [Plasmodium**   3724
         falciparum   AA
         (isolate 3D7)]   <u>align</u>

 Score = 42.7 bits (99), Expect = 0.036
 Identities = 37/172 (21%), Positives = 76/172 (43%), Gaps = 19/172 (11%)

```
Query: 804    LEKENQTDKPSILPQFKRNKAQENSKLDEKVEEPKTSEKVEKEKLSETGNSTSNSTLEEV 863
              +EK+N  +K + +   K+N ++ + +++K + K ++  +K +      N  + S +E
Sbjct: 1099   MEKKNDMEKKNDIE--KKNDMEKKNDMEKKNDMEKKNDMEKKNDMENENNMENKSDIEN- 1155

Query: 864    PTVDPVQEKVAKFAESYGMKLENVLFNMDGTIELYLPSGEVIKKNMADFTGEAPQGNGEN 923
                  + K+ ++ +  MK EN +  + TIE +    +I+ NM +             EN
Sbjct: 1156   ------ENKMNEYKNN--MKNENKMIE-NNTIENNMIENNMIENNMIE------NNMIEN 1200

Query: 924    KPSENGKVSTGTVENQPTENKPADSLPEAPNEKPVKPENSTDNGMLNPEGNV 975
               EN  +   +EN  EN   +    N+  +K E+   N  +N + N+
Sbjct: 1201   NMIENNMIENNMIENNMIENNTKIAYNMNENKNNIKIEDDAKN-QINVQNNI 1251
```

tr <u>Q9L4Z1</u> **Pyruvate dehydrogenase complex subunit E2 [pdhC]**   433
         [Staphylococcus   AA
         epidermidis]   <u>align</u>

```
Score = 42.4 bits (98), Expect = 0.047
Identities = 38/176 (21%), Positives = 71/176 (39%), Gaps = 17/176 (9%)

Query: 782  PVTFMANAYLDNQSTYIVEVPILEKENQTDKPSILPQFKRNKAQENSKLDEKVEEPKTSE 841
            PV+        L ++ T  V   ++ K + D   +   QFK +   E+SK +EK +E      E
Sbjct: 51   PVSGTVEEVLVDEGTVAVVGDVIVKIDAPDAEEM--QFKGHGDDEDSKKEEKEQESPVQE 108

Query: 842  KVEKEKLSETGNSTSNSTLEEVPTVDPVQEKVAKFAESYGMKLENVLFNMDGTIELYLPS 901
            +      E     + T++ +P+         V K+A  G+ ++ V  N  G         +
Sbjct: 109  EASSTQSQEKTEVDESKTVKAMPS-------VRKYARENGVNIKAV--NGSG------KN 153

Query: 902  GEVIKKNMADFTGEAPQGNGENKPSENGKVSTGTVENQPTENKPADSLPEAPNEKP 957
            G + K+++ +           G N ++   S+  V    T+  P     PE   + P
Sbjct: 154  GRITKEDIDAYLNGGSSEEGSNTSAASESTSSDVVNASATQALPEGDFPETTEKIP 209
```

tr Q9VGW4 **CG14692-PA [CG14692] [Drosophila melanogaster (Fruit fly)]** 2762 AA

align

```
Score = 42.4 bits (98), Expect = 0.047
Identities = 49/217 (22%), Positives = 88/217 (39%), Gaps = 12/217 (5%)

Query: 794  QSTYIVEVPILEKENQTDKPSILPQFKRNKAQENSKLDEKVEEPKTSEKVEKEKLSETGN 853
            +ST + E P  EKE  T      L +   ++ + ++ LDEK    K S  ++++   S T
Sbjct: 1131 KSTSLDEKPSSEKEKSTS----LDETPSSEKENSTSLDEKPSPEKESTSLDEKPSSGTEK 1186

Query: 854  STS-----NSTLEEVPTVDPVQEKVAKFAESYGMKLENVLFNMDGTIELYLPSGEVIKKN 908
            STS      +S E+ ++D      + + S  + +   N   +E   P  E        +
Sbjct: 1187 STSLDEKSSSEKEKSTSLDEKPSSEKEKSTSLNERPSSEKENSTSLVENPSPEKESTSLD 1246

Query: 909  MADFTGEAPQGNGENKPSENGKVSTGTVENQPTENKPADSLPEAPNEKPVKPENSTDNGM 968
            +G      + + PS   + ST   E   +E + + S E P++ + K + +
Sbjct: 1247 EKPSSGTEKSTSLDENPSSEKEKSTSLNERPSSEKENSTSQDEKPSSETEKSTSLDEKPS 1306

Query: 969  LNPEGNVGSDPMLDPALEEAPAVD--PVQEKLEKFTA 1003
              E +  D       E++ ++D  P  EK EK T+
Sbjct: 1307 SEKEKSTSLDGKPSSEKEKSTSLDENPSSEK-EKSTS 1342
```

```
Score = 41.2 bits (95), Expect = 0.10
Identities = 56/211 (26%), Positives = 91/211 (42%), Gaps = 38/211 (18%)

Query: 794  QSTYIVEVPILEKENQT---DKPSILPQFKRNKAQENSKLDEKVE-EPKTSEKVEKEKLS 849
            +ST + E P  EKEN T   +KPS         ++ ++++ LDEK   E + S  ++ +   S
Sbjct: 1269 KSTSLNERPSSEKENSTSQDEKPS-------SETEKSTSLDEKPSSEKEKSTSLDGKPSS 1321

Query: 850  ETGNSTSNSTLEEVPTVDPVQEKVAKFAESYGMKLENVLFNMDGTIELYLPSGEVIKKNM 909
            E    STS    L+E P+ +  +EK      E   + EN        T +  PS E   ++
Sbjct: 1322 EKEKSTS---LDENPSSE--KEKSTSLNERPSSEKEN------STSLVENPSPEKESTSL 1370

Query: 910  ADFTGEAPQGNGENKPSENGKVSTGTVENQPTENKPADSLPEAPNEKPVKPENSTDNGML 969
                  + KPS    + ST   EN  +E + +  SL E P+       + ENST
Sbjct: 1371 -----------DEKPSSGTEKSTSLDENPSSEKEKSTSLNERPSS---EKENSTSQDE- 1414

Query: 970  NPEGNVGSDPMLDPALEEAPAVDPVQEKLEK 1000
```

```
                      P          +LD    +     +   V +K+++
Sbjct: 1415 KPSSEKEKSTLLDKNTDLMRDLIQVSQKVDE 1445


 Score = 37.7 bits (86), Expect = 1.2
 Identities = 41/172 (23%), Positives = 75/172 (42%), Gaps = 20/172 (11%)

Query: 804  LEKENQTDKPSILPQFKRNKAQENSKLDEKVEEPKTSEKVEKEKLSETGNSTS------- 856
            LE  + D+ ++L     ++ + ++ LDE    K S ++++  S T  STS
Sbjct: 1068 LEDAGENDRSTLLENLPSSEKENSTSLDENPLPEKESTSLDEKPSSGTEKSTSLDEKSSS 1127

Query: 857  ----NSTLEEVPTVDPVQEKVAKFAESYGMKLENVLFNMDGTIELYLPSGEVIKKNMADF 912
                +++L+E P+ +  +EK   E+  + EN ++D   E   P  E     +
Sbjct: 1128 EKEKSTSLDEKPSSE--KEKSTSLDETPSSEKENST-SLD---EKPSPEKESTSLDEKPS 1181

Query: 913  TGEAPQGNGENKPSENGKVSTGTVENQPTENKPADSLPEAPNEKPVKPENST 964
            +G     ++ + K S   + ST   E   +E + + SL E P+    + ENST
Sbjct: 1182 SGTEKSTSLDEKSSSEKEKSTSLDEKPSSEKEKSTSLNERPSS---EKENST 1230


 Score = 36.6 bits (83), Expect = 2.6
 Identities = 47/183 (25%), Positives = 76/183 (40%), Gaps = 23/183 (12%)

Query: 794  QSTYIVEVPILEKENQT---DKPSILPQ----------FKRNKAQENSKLDEKVEEPKTS 840
            +ST + E P  EKEN T    +KPS   +              R+  Q + K+DE++ + K +
Sbjct: 1394 KSTSLNERPSSEKENSTSQDEKPSSEKEKSTLLDKNTDLMRDLIQVSQKVDEEMSKGKAA 1453

Query: 841  EKVEKEKLSETGNSTSNSTLEEVPTVDPVQEKVAKFAESYGMKLENVLFNMDGTIELYLP 900
                V     G S  NS LE+ P+   +++ A+ +     K    +N+          P
Sbjct: 1454 IAVVDLPDINKGESV-NSPLEKKPSSQELEDIQAELSTD---KETGEPYNLSAEKHKVEP 1509

Query: 901  SGEVIKKNMADFTGEAPQG-NGENKPSENGKVSTGTVENQPTENKPADSLPEAP-NEKPV 958
                 I K   + T     G  G + PS+ GK T   ++   + K D+L  P ++ P
Sbjct: 1510 ----ISKKTIETTDVVDVGLKGNDDPSKPGKTETAATKSVSPDIKDKDTLKLKPDDDSPQ 1565

Query: 959  KPE 961
            K E
Sbjct: 1566 KSE 1568


 Score = 35.8 bits (81), Expect = 4.4
 Identities = 42/180 (23%), Positives = 69/180 (38%), Gaps = 15/180 (8%)

Query: 826  ENSKLDEKVEEPKTSEKVEKEKLSETGNSTSNSTLEEVPTVDPVQEKVAKFAESYGMKLE 885
            EN KL+ ++ + KT  +  +EK S T + + ++          ++   K ES   LE
Sbjct: 1023 ENMKLEGEIGQLKTVFEQSEEKTSPTKSESLHAEDRKIS--------GKSKES----LE 1069

Query: 886  NVLFNMDGTIELYLPSGEVIKKNMADFTG--EAPQGNGENKPSENGKVSTGTVENQPTEN 943
            +   N  T+  LPS E        D     E    + + KPS   + ST   E  +E
Sbjct: 1070 DAGENDRSTLLENLPSSEKENSTSLDENPLPEKESTSLDEKPSSGTEKSTSLDEKSSSEK 1129

Query: 944  KPADSLPEAPNEKPVKPENSTDNGMLNPEGNVGSDPMLDPALEEAPAVDPVQEKLEKFTA 1003
            + + SL E P+ +  K  +      E +   D   P E    +       EK T+
Sbjct: 1130 EKSTSLDEKPSSEKEKSTSLDETPSSEKENSTSLDEKPSPEKESTSLDEKPSSGTEKSTS 1189
```

```
tr    Q90307      Carassius auratus [Carassius auratus (Goldfish)]  798 AA
                                                                    align

  Score = 42.0 bits (97), Expect = 0.061
  Identities = 42/171 (24%), Positives = 63/171 (36%), Gaps = 9/171 (5%)

Query:  831 DEKVEEPKTSEKVEKEKLSETGNSTSNSTLEEVPTVDPVQEKVAKFAESYGMKLENVLFN 890
                +E+ +E    E+ E+ K  E   S      EE  T    +E     AE   G + +
Sbjct:  540 EEEKQEDDAEEEAEEETKAPEAKASPETEKAEEKQTSGGEEE-----AEEEGDEKDEADAG 594

Query:  891 MD--GTIELYLPSGEVIKKNMADFTGEAPQGNGENKPSENGKVSTGTVENQPTENKPADS 948
                 D    E   + E  KK  A+  T E  +G  E +  + K        E +P    KP
Sbjct:  595 SDKGSPGEKEPEAKEQPKKVEAEATKEDKKGKAEKEEPKPEKEKVAVTETKPETAKPESP 654

Query:  949 LPEAPNEKPVKPENSTDNGMLNPEGNVGSDPMLDPALEEAPAVDPVQEKLE 999
                E+P E P    +         E       +P    +P   +EAP  +P  +K E
Sbjct:  655 KSESPKEPP--KTEAPKKEAPKTESPTKEEPKAEPPKKEAPKSEPEDKKKE 703
```

```
tr  Q898B0  Hypothetical protein CTC00555 [CTC00555] [Clostridium tetani]  324 AA

                                                                          align

  Score = 42.0 bits (97), Expect = 0.061
  Identities = 49/168 (29%), Positives = 68/168 (40%), Gaps = 31/168 (18%)

Query:  821 RNKAQENSKL-DEKVEEPKTSEKVEKEKLSETGNSTSNSTLE----EVPTVDPVQEKVAK 875
                + K Q N + + K E PK S+  EK KLSE        L     +P+  P    V
Sbjct:  161 KEKTQPNKHIKNPKKENPKPSQGKEKPKLSEEQMLAKAENLWGGKINLPSYIPKGYDVTN 220

Query:  876 FA--ESYGMKLENVLF---NMDGTIELYLPSGEVIKKNMADFTGEAPQGNGENKPSENGK 930
                  +YG K+  + +    N   +EL +  G+   K    D      GNG   P EN K
Sbjct:  221 IGMDTAYGSKVLKITYKNPNSGNFLELKVLEGD--KTAFED-------GNGSKNPKENEK 271

Query:  931 VSTGTVENQPTEN--KPADSLPEAPNEKPVKPENSTDNGM-LNPEGNV 975
                ++  P EN   KP ++          VK  NS    NG+  N EGN+
Sbjct:  272 NEDDNRQDNPKENVDKPTET---------VKSVNSNKNGVEYNIEGNI 310
```

```
tr  Q6R4Z8  Dehydrin cor29 [Capsella bursa-pastoris (Shepherd's purse)]  261 AA

                                                                          align

  Score = 42.0 bits (97), Expect = 0.061
  Identities = 31/113 (27%), Positives = 55/113 (48%), Gaps = 18/113 (15%)

Query:  792 DNQSTYIVEVPILEKENQTDKPSILPQFKRNKAQENSKLDEKVEEPKTSEKVE------- 844
                +N+ T + E+    +E++ +KPS++ + + R+ +  +S  DE+ EE K  + VE
Sbjct:   80 ENKITLLEELQEKTEEDEENKPSVIEKLHRSNSSSSSSSDEEGEEKKKKKTVEGEEEKKG 139

Query:  845 -----KEKLSETGNSTSNSTLEEVPTVD----PVQEKVAKFAESYGMKLENVL 888
                  KEKL   G+   +  +VP V      PV E V  +  E+ G + + V+
Sbjct:  140 AMDKIKEKL--PGHHDKETEDHDVPVVSTIQVPVSESVVEHHETEGEEKKGVM 190
```

```
tr  Q8MMQ1  Similar to Required for the transfer of mannosylphosphate to   393
            cell                                                           AA
            wall mannans.; Mnn4p [Dictyostelium discoideum (Slime          align
            mold)]

 Score = 41.6 bits (96), Expect = 0.080
 Identities = 40/171 (23%), Positives = 76/171 (44%), Gaps = 15/171 (8%)


Query:  805  EKENQTDKPSILPQFKRNKAQENSKLDEKVEEPKTSEKVEKEKLSETGNST--------S 856
             E+E + +K      P+     K + ++K +E+ E+ +   EKVE+E+ SET +T        +
Sbjct:  205  EEEEEEEKVEEKPKKVIKKQKVSNKTEEQKEKVEEEEKVEEEEQSETTTTTQKKKKFISN 264

Query:  857  NSTLEEVPTVDPVQEKVAKFAESYGMKLENVLFNMDGTIELYLPSGEVIKKN---MADFT 913
             +ST++ + +   QEK A   + K     L + + +E      V++K   +
Sbjct:  265  DSTIDIMKQISEEQEKKANQKITQEKKKTQQLQSKE--VEKQTKKKNVVEKKKTILKKLK 322

Query:  914  GEAPQGNGENKPSENGKVSTGTVENQPTENKPADSLPEAPNEKPVKPENST 964
             GE Q   + K E   V+ T   + + +K   + E P+E+   P   +
Sbjct:  323  GEEEQ--EQEKEEEKPVVAKTTTTTKKSVSKKPKTQVEQPSEEQTTPSKKS 371
```

```
tr  Q8IBL1  Hypothetical protein MAL7P1.129 [MAL7P1.129] [Plasmodium       1003
            falciparum                                                     AA
            (isolate 3D7)]                                                 align

 Score = 41.6 bits (96), Expect = 0.080
 Identities = 53/233 (22%), Positives = 87/233 (36%), Gaps = 32/233 (13%)

Query:  402  YIPKSNQIGQPTLPNNSLATPSPSLPINPGISHEKHEEDGYGFDANRII--AEDESGFIM 459
             Y P   + +LP+N    S S +NP   +K        +N I   EDE
Sbjct:  338  YSPNKTKNNNSSLPDNKKTNKSNSY-VNPTQKKDKINNSNTTVVSNNISHEKEDEKRKKS 396

Query:  460  SHGNHNHYFFKKDLTEEQIKAAQKHLEEVKTSHNGLDSLSSHEQDYPGNAKEMKDLDKKI 519
             H + N    KK++  +  K + + +   K  +NG +  ++H    N         +KK+
Sbjct:  397  DHRDLNETE-KKNIKLDYPKDFESNNKHDKHINNGDEEKNAHSSHSQNNK------EKKL 449

Query:  520  EEKIAGIMKQYGVKRESIVVNKE----------------KNAIIYPHGDHHHADPIDEH 562
             +++ +KR ++  K+             K I+   + H++ I+EH
Sbjct:  450  HSNDTEALRRSSIKRSTVDYEKKNLKSESYYGKEYLNSYSKEDIMKEGKEKKHSNDINEH 509

Query:  563  KPVGIGHSHSN--YELFKPEEGVAKKEGNKVYTGEELTNVVNLLKNSTFNNQN 613
             +    SHSN  E  K E+   KK+ N          NV    K  T NN N
Sbjct:  510  NNIEGSLSHSNRLRESIKFED---KKKNNNSNNYSHYNNVHGAFKKHTHNNDN 559
```

```
sp  Q8CT13    Dihydrolipoyllysine-residue acetyltransferase component of  433
    ODP2_STAEP  pyruvate                                                  AA
                dehydrogenase complex (EC 2.3.1.12) (E2)                   align
                (Dihydrolipoamide acetyltransferase component of
                pyruvate dehydrogenase complex) [pdhC] [Staphylococcus
                epidermidis]

 Score = 41.2 bits (95), Expect = 0.10
 Identities = 38/176 (21%), Positives = 70/176 (39%), Gaps = 17/176 (9%)
```

```
Query:  782  PVTFMANAYLDNQSTYIVEVPILEKENQTDKPSILPQFKRNKAQENSKLDEKVEEPKTSE  841
             PV+       L ++ T  V   ++ K +  D    +  QFK +    E+SK +EK +E      E
Sbjct:   51  PVSGTVEEVLVDEGTVAVVGDVIVKIDAPDAEEM--QFKGHGDDEDSKKEEKEQESPVQE  108


Query:  842  KVEKEKLSETGNSTSNSTLEEVPTVDPVQEKVAKFAESYGMKLENVLFNMDGTIELYLPS  901
             +   +   E    +  T++ +P+        V K+A   G+ ++ V  N  G          +
Sbjct:  109  EASSTQSQEKTEVDESKTVKAMPS-------VRKYARENGVNIKAV--NGSG------KN  153


Query:  902  GEVIKKNMADFTGEAPQGNGENKPSENGKVSTGTVENQPTENKPADSLPEAPNEKP  957
             G + K+++ +         G N   +   S+ V   T+ P   PE   + P
Sbjct:  154  GRITKEDIDAYLNGGSSEEGSNTSVASESTSSDVVNASATQALPEGDFPETTEKIP  209
```

tr   O33741   **SpsA protein [Streptococcus pneumoniae]**   539 AA
                                                                align

 Score = 41.2 bits (95), Expect = 0.10
 Identities = 48/197 (24%), Positives = 82/197 (41%), Gaps = 21/197 (10%)

```
Query:  794  QSTYIVEVPILEKENQTDKPSILP--------QFKRNKAQENSKLDE---KVEEPKTSEK  842
             ++ Y+ E+ +LE++++ + PS +        +FK++  +   K+ E    KVEE  K   +
Sbjct:   97  KTKYLRELNVLEEKSKDELPSEIKAKLDAAFEKFKKDTLKPGEKVAEAKKKVEEAKKKAE  156


Query:  843  VEKEKLSETGNSTSNSTLE-EVPTVDPVQEKVAKFAESYGMKLENVLFNMDGTIELYLPS  901
             +KE+        + + TLE E+    D      K AE   +K E    F  +GTI+       +
Sbjct:  157  DQKEEDRRNYPTNTYKTLELEIAEFDVK----VKEAELELVKEEAKEFRNEGTIK---QA  209


Query:  902  GEVIKKNMADFTGEAPQGNGENKPSENGKVSTGTVENQPTENKPADSLPEAPNEKPVKPE  961
             E ++    A+ T           K  E K      E   + KPA+    AP  +P KP
Sbjct:  210  KEKVESKKAEATRLENIKTDRKKAEEEAKRKAA--EEDKVKEKPAEQPQPAPATQPEKPA  267


Query:  962  NSTDNGMLNPEGNVGSD  978
                      +      P+      D
Sbjct:  268  PKPEKPAEQPKAEKTDD  284
```

tr Q8I1W5 **Hypothetical protein PFD0320c (Fragment) [PFD0320c]**   2228
            [Plasmodium   AA
            falciparum (isolate 3D7)]   align

 Score = 41.2 bits (95), Expect = 0.10
 Identities = 37/161 (22%), Positives = 60/161 (36%), Gaps = 6/161 (3%)

```
Query:  805  EKENQTDKPSILPQFKRNKAQENSKLDEKVEEPKTSEKVEKEKLSETGNSTSNSTLEEVP  864
             EKEN T +   + + K N +E K+ E+ E+       EK + + E      +    + EE
Sbjct: 1370  EKENMTAEKENMTEEKENMTEEKEKMTEEKEQVTEKEKEKMTEEKENMTEKKENMTEEKE  1429


Query:  865  TVDPVQEKVAKFAESYGMKLENVLFNMDGTIELYLPSGEVIKKNMADFTGEAPQGNGENK  924
             +   +E++ + E    + E +   + E    + E   I +   T E  Q   E +
Sbjct: 1430  QITEEKEQITEEKEQMTEEKEQITEEKEQITE----EKEQITEEKEQITEEKEQITEEKE  1485


Query:  925  PSENGKVSTGTVENQPTENKPADSLPEAPNEKPVKPENSTD  965
                  K       + Q TE K + + E       K EN TD
Sbjct: 1486  QITEEKEQITEEKEQITEEK--EQITEEKENMTEKKENMTD  1524
```

tr   Q07593   **K2 protein (Fragment) [Entamoeba histolytica]**   238 AA
                                                                      align

 Score = 41.2 bits (95), Expect = 0.10
 Identities = 46/205 (22%), Positives = 79/205 (38%), Gaps = 18/205 (8%)

```
Query: 803  ILEKENQTDKPSILPQFKRNKAQENSKLDEKVEEPKTSEKVEKEKLSETGNSTSNST--- 859
            IL+ + +    I   F +N+A      EK+EE K  E+  EK   S TG+   N++
Sbjct: 14   ILDIDQEVKDTKIYSVFLKNEASP-----EKLEENKEDEEKEKSSSSNTGDEPDNNSEAK 68

Query: 860  LEEVPTVDPVQEKVAKFAESYGMKLENVLFNMDGTIELYLPSGEVIKKNMADFTGEAPQG 919
            ++ P    +  K  S  K E      + +    S +        + + P+
Sbjct: 69   SDDEPEASSNDKSDDKVEASSSDKPE-----ASSSDKPEASSSDKPDNKPEASSSDKPEE 123

Query: 920  NGENKPSENGKVSTGTVENQPTENKPADSLPEAPNEKPVKPENSTDNGMLNPEGNVGSDP 979
            +  +KP +        +     +NKP +S + P+ KP    +S+D    PE +      P
Sbjct: 124  SSSDKPEASSSDKPEASSSDKPDNKPEESSSDKPDNKP--EASSSDKPDNKPEASSSDKP 181

Query: 980  MLDPALEEAPAVDPVQEKLEKFTAS 1004
               P  EA + D   K E +A+
Sbjct: 182  DNKP---EASSSDKPDNKPEASSAT 203
```


sp   Q28820   **Triadin [TRDN] [Oryctolagus cuniculus**   705 AA
      TRDN_RABIT   **(Rabbit)]**   align

 Score = 40.8 bits (94), Expect = 0.14
 Identities = 42/186 (22%), Positives = 68/186 (35%), Gaps = 12/186 (6%)

```
Query: 806  KENQTDKPSILPQFKRNKAQENSKLDEKVEEPKTSEKVEKEKLSETGNSTSNSTLEEVPT 865
            K+   ++P   Q K+ K +E +K  +K E   SEK  K+     ++ ST + VP
Sbjct: 384  KKPAEEQPKGKKQEKKEKHEEPAKSTKK-EHAAPSEKQAKAKIERKEEVSAASTKKAVPA 442

Query: 866  VDPVQEKVAKFAESYGMKLENVLFNMDGTIELYLPSGEVIKKNMADFTGEAPQGNGENKP 925
            +EK  K  E    K        G I  L  E+ K+        +   E K
Sbjct: 443  KK--EEKTTKTVEQETRK------EKPGKISSVLKDKELTKEKEVKVPASLKEKGSETKK 494

Query: 926  SENGKVSTGTVENQPTENKPADSLPEAPNEKPVKPENSTDNGMLNPEGNVGSDPMLDPAL 985
            E         ++    E KP   +  P + +K E   +  ++ PE        +   L
Sbjct: 495  DEKTSKPEPQIKK---EEKPGKEVKPKPPQPQIKKEEKPEQDIMKPEKTALHGKPEEKVL 551

Query: 986  EEAPAV 991
            ++  AV
Sbjct: 552  KQVKAV 557
```

 Score = 35.8 bits (81), Expect = 4.4
 Identities = 22/82 (26%), Positives = 40/82 (47%), Gaps = 2/82 (2%)

```
Query: 799  VEVPILEKENQTDKPSILPQFKRNKAQENSKLDEKVEEPKTSEKVEKEKLSETGNSTSNS 858
            +    ++ KE + +K  + + K    + + K  K  + + EK E+P+T   ++EK + T
Sbjct: 158  IPTKVVHKEKEKEKEKVKEKEKPEKKATHKEKLEKKEKPETKTVTKEEKKARTKEKIEEK 217

Query: 859  TLEEVPTV--DPVQEKVAKFAE 878
            T +EV  V  + V++ VAK E
```

Sbjct: 218 TKKEVKGVKQEKVKQTVAKAKE 239

tr Q7SXW9 **Hypothetical protein (Fragment) [Brachydanio rerio (Zebrafish)** 232 AA
       (Danio rerio)]
                                                                        align

 Score = 40.8 bits (94), Expect = 0.14
 Identities = 37/134 (27%), Positives = 55/134 (40%), Gaps = 18/134 (13%)

Query: 826 ENSKLDEKVEEPKTSEKVEKEKLSETGNSTSNSTLEEVPTVDPVQEKVAKFAESYGMKLE 885
           +N DE E K +E K   E G++ +N   P E A+ ++ G
Sbjct: 12 KNGAKDETAAE-KPAEAANKSNGQENGHAKTNGN------ASPNAEAAAEDVQANGKHAA 64

Query: 886 NVLFNMDGTIELYLPSGEVIKKNMADFTGEAPQGNGENKPSENGKVSTGTVENQPTENKP 945
                   DG ++   E K   AD   AP+G GE+   NG+ ST T E+ T ++P
Sbjct: 65 ------DGEVK-----AEEGKAEEADAEKAAPEGEGESSAVANGEDSTKTEESAATSSEP 113

Query: 946 ADSLPEAPNEKPVK 959
           A +   +KP K
Sbjct: 114 AKTKKRFSFKKPFK 127

tr Q8DPR5 **Immunoglobulin A1 protease (EC 3.4.24.13) [iga]**   1963
       [Streptococcus   AA
       pneumoniae (strain ATCC BAA-255 / R6)]   align

 Score = 40.8 bits (94), Expect = 0.14
 Identities = 45/183 (24%), Positives = 74/183 (39%), Gaps = 27/183 (14%)

Query: 800 EVPILEKENQTDKPSILP-QFKRNKAQENSKLDEKVEEPKTSEKVEKEKLSETGNSTSNS 858
           E P+ E T+ SI +   A+E++ EKV +SE   + + ++TS
Sbjct: 512 ETPVNPNEGTTEGTSIQEAAENPVQPAEESTTNSEKVSPDTSSENTGEVSSNPSDSTTSVG 571

Query: 859 TLEEVPTVDPVQEKVAKFAESYGMKLENVLFNMDGTIELYLPSGEVIKKNMADFTGEAPQ 918
             +   D E K E +   +GT+E   ++   E P
Sbjct: 572 ESNKPEHNDSKNENSEKTVEEVPVNPN------EGTVE-----------GTSNQETEKPV 614

Query: 919 GNGENKPSENGKVS---TGTVENQPTENKP--ADSLPEAPNEKPVKPENS----TDNGML 969
           E + +GK++   TG V N+P+++KP +S   N   KPENS   ++NG
Sbjct: 615 QPAEETQTNSGKIANENTGEVSNKPSDSKPPVEESNQPEKNGTATKPENSGNTTSENGQT 674

Query: 970 NPE 972
           PE
Sbjct: 675 EPE 677

tr Q59947 **Immunoglobulin A1 protease (EC 3.4.24.13) [iga]**   1964
       [Streptococcus   AA
       pneumoniae]   align

 Score = 40.8 bits (94), Expect = 0.14
 Identities = 45/183 (24%), Positives = 74/183 (39%), Gaps = 27/183 (14%)

```
Query:  800  EVPILEKENQTDKPSILP-QFKRNKAQENSKLDEKVEEPKTSEKVEKEKLSETGNSTSNS 858
             E P+    E  T+   SI   +     A+E++   EKV    +SE    +   + ++TS
Sbjct:  512  ETPVNPNEGTTEGTSIQEAENPVQPAEESTTNSEKVSPDTSSENTGEVSSNPSDSTTSVG 571

Query:  859  TLEEVPTVDPVQEKVAKFAESYGMKLENVLFNMDGTIELYLPSGEVIKKNMADFTGEAPQ 918
                +    D  E   K   E    +       +GT+E           ++    E P
Sbjct:  572  ESNKPEHNDSKNENSEKTVEEVPVNPN------EGTVE-----------GTSNQETEKPV 614

Query:  919  GNGENKPSENGKVS---TGTVENQPTENKP--ADSLPEAPNEKPVKPENS----TDNGML 969
                   E  + +GK++    TG V N+P+++KP   +S      N     KPENS    ++NG
Sbjct:  615  QPAEETQTNSGKIANENTGEVSNKPSDSKPPVEESNQPEKNGTATKPENSGNTTSENGQT 674

Query:  970  NPE 972
             PE
Sbjct:  675  EPE 677
```

tr Q28688 **Neurofilament-H (Fragment) [Oryctolagus cuniculus (Rabbit)]** 511 AA

align

```
 Score = 40.8 bits (94), Expect = 0.14
 Identities = 54/233 (23%), Positives = 92/233 (39%), Gaps = 18/233 (7%)

Query:  780  KIPVTFMANAYLDNQSTYIVEVPILEKENQTDKPSILPQFKRNKAQENSKLDEKVEEPKT 839
             K PV    A +     +S   + P+ E+    +K     P +   K+ E +K  EK + P
Sbjct:  275  KSPVKEEAKSPEKAKSPEKAKSPVKEEAKSPEKAKS-PVKEEAKSPEKAKSPEKAKSPVK 333

Query:  840  SEKVEKEKLSETGNSTSNSTLEEVPTVD----PVQEKVAKFAESYGMKLENV-LFNMDGT 894
             E    EK +++       +    EE  + +     PV+E+ AK  E    K EV +
Sbjct:  334  EEAKSPEK-AKSPEKAKSPVKEEAKSPEKAKSPVKEE-AKSPEKETPKKEEVKVKEPPKK 391

Query:  895  IELYLPSGEVIKKNMADFTGEAPQGNGENKPSENGKVSTGTVENQPTENKPADSLPEAPN 954
             +E   P+   ++K+       EAP+       E  K ST   +   E+K   +   +
Sbjct:  392  VEETAPAPPKVEKDSKK--DEAPKKEAPKPAVEKPKESTAEAKKDEAEDKKTAAPAKVEG 449

Query:  955  EKPVKPENSTDNGMLNPEGNVGSDPML-------DPALEEAPAVDPVQEKLEK 1000
             ++  KP+ T+       PE     +P      P  EE PA  PV++  ++
Sbjct:  450  KEEAKPKEKTEVAKKEPEDAKAKEPSKPTEKEPEKPKKEETPAA-PVKKNAKE 501
```

tr Q94674 **Thrombospondin-related anonymous protein (Fragment) [TRAP]** 614 AA
          [Plasmodium gallinaceum]

align

```
 Score = 40.8 bits (94), Expect = 0.14
 Identities = 44/209 (21%), Positives = 84/209 (40%), Gaps = 28/209 (13%)

Query:  792  DNQSTYIVEVPILEKENQTDKPSILPQFKRNKAQENSKLDEKVEEPKTSEKVEKEKLSET 851
             D +   I E    E   +  KP   P+ K  ++     K  E VEK  EV +EK  E+
Sbjct:  299  DEEPEPIPEEKKPEPVPEEKKPESAPEEKNPESVPEEKKPESVPEEKEPESVPEEKEPES 358

Query:  852  ---GNSTSNSTLEEVPTVDPVQEKVAKFAESYGMKLENVLFNMDGTIELYLPSGEVIKKN 908
                ++  E+ P  DP ++K+    E   G  K+E +              P   E ++
```

```
Sbjct:  359  VPEEKEPESAPEEKKPESDPEEKKLEPIPE--GKKIEPI-----------PEEEKLEPI  404

Query:  909  MADFTGEAPQGNGENKPSENGKVSTGTVENQPTENKPADSLPEAPNEKPVKPENSTDN--  966
                 +    E+   + E++P  +G+ +    +N P + +    + PN++ + P+N  D+
Sbjct:  405  PEEKKPESVTEDRESEPVPDGE-AENVPQNIPDDEQEEKISGDIPNDEELIPKNEPDDIK  463


Query:  967  --------GMLNPEGNVGSDPMLDPALEE  987
                     ++  P+     + + +P  EE
Sbjct:  464  RNEYDTTPNIIPPKDTYNDNEITNPISEE  492
```

tr Q6BLN0 **Similar to ca|CA2433|IPF12959 Candida albicans IPF12959**   782
                unknown   AA
                function [DEHA0F13211g] [Debaryomyces hansenii CBS767]   align

 Score = 40.8 bits (94), Expect = 0.14
 Identities = 40/213 (18%), Positives = 90/213 (41%), Gaps = 10/213 (4%)

```
Query:  800  EVPILEKENQTDKPSILPQFKRNKAQENSKLDEKVEEPKTSEKVEKEKLSETGNSTSNST  859
             EV  L  E  T+ P +  + ++ + QE    + + +  + E  E   +S+  +S   +
Sbjct:  315  EVKRLPGEVDTELPIVKQELEQKRIQE----ERESKSNQIEETAETTAVSQNDDSNEENK  370

Query:  860  LEEVPTVDPVQEKVAKFAESYGMKLE---NVLFNMDGTIELYLPSGEVIKKNMA-DFTGE  915
               E+  + P ++ ++ +E   + E    + N +   +     N++ + T +
Sbjct:  371  NIEIDNITPPSKESSELSEQVDVSKEVTGESISNEEVETQTITEDDKNANPNISREATIK  430

Query:  916  APQGNGENKPSENGKVSTGTVENQPTENKPADSLPEA-PN-EKPVKPENSTDNGMLNPEG  973
               +  N+ SEN +V   ++N  +    DS+  + PN + P+K E S+++ ++N
Sbjct:  431  ENTASNSNQASENDEVIDNDLDNNDEKLTENDSVSNSKPNVDAPIKEEQSSNDDLINESS  490

Query:  974  NVGSDPMLDPALEEAPAVDPVQEKLEKFTASYG  1006
               ++   D  E  A+D +   E    + G
Sbjct:  491  KEDTNKPKDDDGETQAAIDSIVSDEENLLNNKG  523
```

sp_vs Q28820- **Splice isoform Cardiac 3 of Q28820 [TRDN] [Oryctolagus**   668
          4   **cuniculus**   AA
                      (Rabbit)]   align

 Score = 40.8 bits (94), Expect = 0.14
 Identities = 42/186 (22%), Positives = 68/186 (35%), Gaps = 12/186 (6%)

```
Query:  806  KENQTDKPSILPQFKRNKAQENSKLDEKVEEPKTSEKVEKEKLSETGNSTSNSTLEEVPT  865
             K+  ++P    Q K+ K +E +K  +K E    SEK  K K+    ++ ST + VP
Sbjct:  384  KKPAEEQPKGKKQEKKEKHEEPAKSTKK-EHAAPSEKQAKAKIERKEEVSAASTKKAVPA  442

Query:  866  VDPVQEKVAKFAESYGMKLENVLFNMDGTIELYLPSGEVIKKNMADFTGEAPQGNGENKP  925
               +EK  K E    K       G I  L   E+ K+       +   E K
Sbjct:  443  KK--EEKTTKTVEQETRK------EKPGKISSVLKDKELTKEKEVKVPASLKEKGSETKK  494

Query:  926  SENGKVSTGTVENQPTENKPADSLPEAPNEKPVKPENSTDNGMLNPEGNVGSDPMLDPAL  985
             E         ++   E KP  +   P +  +K E   + ++ PE        + L
Sbjct:  495  DEKTSKPEPQIKK---EEKPGKEVKPKPPQPQIKKEEKPEQDIMKPEKTALHGKPEEKVL  551

Query:  986  EEAPAV  991
```

```
                ++  AV
Sbjct: 552 KQVKAV 557
```


```
 Score = 35.8 bits (81), Expect = 4.4
 Identities = 22/82 (26%), Positives = 40/82 (47%), Gaps = 2/82 (2%)

Query: 799 VEVPILEKENQTDKPSILPQFKRNKAQENSKLDEKVEEPKTSEKVEKEKLSETGNSTSNS 858
            +    ++ KE + +K  +  + K  K    + +  EK E+P+T    ++EK + T
Sbjct: 158 IPTKVVHKEKEKEKEKVKEKEKPEKKATHKEKLEKKEKPETKTVTKEEKKARTKEKIEEK 217

Query: 859 TLEEVPTV--DPVQEKVAKFAE 878
            T +EV  V  + V++ VAK  E
Sbjct: 218 TKKEVKGVKQEKVKQTVAKAKE 239
```


```
sp_vs  Q28820-    Splice isoform Skeletal 3 of Q28820 [TRDN] [Oryctolagus    698
       6           cuniculus                                                AA
                   (Rabbit)]                                                align
```

```
 Score = 40.8 bits (94), Expect = 0.14
 Identities = 42/186 (22%), Positives = 68/186 (35%), Gaps = 12/186 (6%)

Query: 806 KENQTDKPSILPQFKRNKAQENSKLDEKVEEPKTSEKVEKEKLSETGNSTSNSTLEEVPT 865
            K+   ++P   Q K+ K +E +K  +K E     SEK  K K+       ++ ST + VP
Sbjct: 384 KKPAEEQPKGKKQEKKEKHEEPAKSTKK-EHAAPSEKQAKAKIERKEEVSAASTKKAVPA 442

Query: 866 VDPVQEKVAKFAESYGMKLENVLFNMDGTIELYLPSGEVIKKNMADFTGEAPQGNGENKP 925
            +EK  K  E       K     G I    L   E+ K+        +     E K
Sbjct: 443 KK--EEKTTKTVEQETRK------EKPGKISSVLKDKELTKEKEVKVPASLKEKGSETKK 494

Query: 926 SENGKVSTGTVENQPTENKPADSLPEAPNEKPVKPENSTDNGMLNPEGNVGSDPMLDPAL 985
            E        ++   E KP  +   P +  +K E   +  ++ PE          +  L
Sbjct: 495 DEKTSKPEPQIKK---EEKPGKEVKPKPPQPQIKKEEKPEQDIMKPEKTALHGKPEEKVL 551

Query: 986 EEAPAV 991
            ++  AV
Sbjct: 552 KQVKAV 557
```


```
 Score = 35.8 bits (81), Expect = 4.4
 Identities = 22/82 (26%), Positives = 40/82 (47%), Gaps = 2/82 (2%)

Query: 799 VEVPILEKENQTDKPSILPQFKRNKAQENSKLDEKVEEPKTSEKVEKEKLSETGNSTSNS 858
            +    ++ KE + +K  +  + K  K    + +  EK E+P+T    ++EK + T
Sbjct: 158 IPTKVVHKEKEKEKEKVKEKEKPEKKATHKEKLEKKEKPETKTVTKEEKKARTKEKIEEK 217

Query: 859 TLEEVPTV--DPVQEKVAKFAE 878
            T +EV  V  + V++ VAK  E
Sbjct: 218 TKKEVKGVKQEKVKQTVAKAKE 239
```


```
tr    Q54875      IgA1 protease [iga] [Streptococcus pneumoniae]   1927 AA
                                                                    align
```

```
Score = 40.4 bits (93), Expect = 0.18
Identities = 67/265 (25%), Positives = 109/265 (40%), Gaps = 43/265 (16%)

Query:  792  DNQSTYIVEV---PILEKENQTDK--PSILPQFKRNKAQENSK--LDEKVEEPKTSEKVE  844
             D QS IVE    P L+  +DK  P + P         +  +  + ++ E   S+K E
Sbjct:  409  DVQSGAIVEPAIQPELPEAVVSDKGEPEVQPTLPEAVVTDKGEPAVQPELPEAVVSDKGE  468

Query:  845  KEKLSETGNSTSNSTLEEVPTVDPVQEKVAKFAESYGMKLENVLFNMDGTIELYLPSGEV  904
             E+++      N  +E+V   PV+    K E    KEV   +  TE  +    E
Sbjct:  469  PEQVAPLPEYKGN--IEQVKPETPVE----KTKEQGPEKTEEV--PVKPTEETPVNPNEG  520

Query:  905  IKKNMADFTGEAPQGNGENKPSENGKVS---TGTVENQPTENKP--ADSLPEAPNEKPVK  959
             + +   E P   E+  + +GK++    TG V N+P+++KP   +S   N    K
Sbjct:  521  TTEGTSIQGAENPVQPAEDTQTNSGKIANENTGEVSNKPSDSKPPVEESNQPEKNGTATK  580

Query:  960  PENS----TDNGMLNPE---GNVGSDPMLDPALEEAPAVDPVQEKLEKFTASYGLGLDSV  1012
             PENS   ++NG  PE   GN   D       +   +  ++++ E      L   D
Sbjct:  581  PENSGNTTSENGQTEPEPSNGNSTEDVSTKSNTSNSNGNEEIKQENE-------LDPDKK  633

Query:  1013 IFNMDGTIELRLPSGEVIKKNLSDL  1037
             + + + T+ELR       N+SDL
Sbjct:  634  VEDPEKTLELR---------NVSDL  649
```

```
tr   Q9GUY4    Crustocalcin [DD4(ccn)] [Penaeus japonicus (Kuruma    831 AA
                prawn)]                                                align

 Score = 40.4 bits (93), Expect = 0.18
 Identities = 54/241 (22%), Positives = 90/241 (36%), Gaps = 22/241 (9%)

Query:  805  EKENQTDKPSILPQFKRNKAQENSKLDEKVEEPKTSEKVEKEKLSETGNSTSNSTLEEVP  864
             E E +++      + +  + + S +E+ EEP+     +EK +E G       EE
Sbjct:  119  ESEGRSEGTQEEEEEEEEEGESESSGNERSEEPEKEGAEGEEKSAEPG-------AEEAE  171

Query:  865  TVDPVQEKVAKFAESYGMKLENVLFNMDGTIELYLPSGEVIKKNMADFTGE---APQGNG  921
             +   E  AK AE+ G   E     D  E      GE  K  ++   +   A +
Sbjct:  172  GEEEAAEPEAKGAEAEGKSAETEKGKSDAEAE-EEAEGEGASKKASEENADVEKAQEVEA  230

Query:  922  ENKPSENGKVSTGTVENQPTENKPADSLPEAPNEKPVKPENSTDNGMLNPEGNV------  975
             + +  + + G      E+   PE P   PVK E++      E  V
Sbjct:  231  ASSSVDEVAAAEGDASEAAAESTAKQLSPETPALVPVKVESTFAPVSPVTEAPVIATAAM  290

Query:  976  --GSDPMLDPALEEAPAVD-PVQEKLEKFTASYGLGLDSVIFNMDGTI--ELRLPSGEVI  1030
              + P++ P   E PA   P E     +AS GLG  ++  +   E+ LP+ EV+
Sbjct:  291  LPAAAPLVAPVTVETPAAPVPSPEAYSADSASGGLGSQDLVLVVIPPVEQEISLPNAEVV  350

Query:  1031 K  1031
             +
Sbjct:  351  E  351
```

```
tr   Q9GTX2    Glutamate-rich protein [GLURP] [Plasmodium            1236 AA
                falciparum]                                            align

 Score = 40.4 bits (93), Expect = 0.18
```

Identities = 49/185 (26%), Positives = 84/185 (44%), Gaps = 27/185 (14%)

```
Query: 791  LDNQSTYIVE----VPILEKENQTDKPSILPQFKRN-KAQENSKLDEKVEEPKTSEKVEK 845
            LDNQ +I +     + +L+ EN +    + PQ K N ++ E  +D ++  P   E VE
Sbjct: 361  LDNQKEHIDQSQHNINVLQ-ENNINNHQLEPQEKPNIESFEPKNIDSEIILP---ENVET 416

Query: 846  EKLSETGNSTSNS---TLEEVPTVDPVQEKVAKFAESYGMKLENVLFNMDGTIELYLPSG 902
            E++ +   S +S    T EE +    +E V++                N    T+E
Sbjct: 417  EEIIDDVPSPKHSNHETFEEETSESEHEEAVSEK------------NAHETVEHEETVS 463

Query: 903  EVIKKNMADFTGEAPQGNGENKPSENGKVSTGTVENQPTENKPADSLPEAPNEKPVKPEN 962
            +       AD G   Q N  N+ +EN V +    E++P EN+ +  SL E  +E+ V  +N
Sbjct: 464  QESNPEKADNDGNVSQ-NSNNELNENEFVESEKSEHEPAENEES-SLEEGHHEEIVPEQN 521

Query: 963  STDNG 967
            +  ++G
Sbjct: 522  NEESG 526
```

tr Q8IJ56 **Glutamate-rich protein [PF10_0344] [Plasmodium falciparum**   1233
           (isolate   AA
           3D7)]   align

  Score = 40.4 bits (93), Expect = 0.18
  Identities = 49/185 (26%), Positives = 84/185 (44%), Gaps = 27/185 (14%)

tr Q6FWC0 **Candida glabrata strain CBS138 chromosome D complete sequence** 1148 AA
           [CAGL0D01364g] [Candida glabrata CBS138]
                                                                                align

  Score = 40.4 bits (93), Expect = 0.18
  Identities = 58/253 (22%), Positives = 107/253 (41%), Gaps = 40/253 (15%)

```
Query: 753  LENNYKVGEIKLPIPKL-----NQGTTRTAGNKIPVTFMANAYLDNQSTYIVE------V 801
            +EN +    LP+P       N T  TA + +   M +A D +    VE     +
Sbjct: 783  VENGQEETTTSLPVPATEETNHNNETVGTAED-LKDEAMHDATQDGEKHQAVEESVTTDL 841

Query: 802  PILEKENQTDKPSILPQFKRNKAQENSKLDEK-VEEPKTSEKVEKEKLSETGNSTSNSTL 860
```

```
                 P  E+E   D+P+      ++++  +EN++ +     EEP  S+  +  +K SET  +T  +
Sbjct:  842  PSTEQEKHLDEPA-----QKSEVEENTRAETAGYEEP--SDLPDAQKSSETTGTTETAEA 894

Query:  861  EEVPTVDPVQE----KVAKFAESYGMKLENVLFNMDGTIELYLPSGEVIKKNMADFTGEA 916
             E    +   E   + AK AE+   +L  + N +  IE   + E + +    E
Sbjct:  895  EHTEVTNVPSEVETSETAKPAETTADQLTGEI-NQEKQIEQPSEASESVPEEKTEV--EQ 951

Query:  917  PQGNGENKPSENGKVSTGTVENQPTENKPADSLPEAPNEKPVKPENSTDNGMLNPEGNV- 975
             P       KP+E      T + + + +   ++ +  + +PVK  +ST+    +P  N+
Sbjct:  952  P------KPAEEDSKGNDTPQEEVSISNTETTMADESSTQPVK--DSTETNQNDPADNMT 1003

Query:  976  ----GSDPMLDPA 984
                 G D  +D A
Sbjct: 1004  EESHGQDSTIDEA 1016
```

tr Q6CTI0 **Similar to sp|Q05050 Saccharomyces cerevisiae YMR031c**   830
           [KLLA0C12573g]   AA
           [Kluyveromyces lactis NRRL Y-1140]   align

 Score = 40.4 bits (93), Expect = 0.18
 Identities = 48/219 (21%), Positives = 83/219 (36%), Gaps = 17/219 (7%)

```
Query:  799  VEVPILEKENQTDKPSILPQFKRNKAQENSKLDEKVEEPKTSEKVEKEKLSETGNSTSNS 858
             +EV +  E QT+ +   + KR K +    + E V E K  +K+++EK       +TS
Sbjct:  593  LEVDAIINERQTELSNTEIEVKREKLKLIDAMKE-VAEVKNEDKIDEEKAKAFLGTTSGE 651

Query:  859  TLEEVPTVDPVQ--EKVAKFAE--SYGMKLENVLFNMDGTIELYLPSGEVIKKNMADFTG 914
             L     V+P  +    K +E S   K+E V N+   +    P+ + K ++   T
Sbjct:  652  FLASQKKVEPATKLQSDPKLSEPSSKSTKIEGVTGNVKADVPASPPAHK--KHSIGGLTS 709

Query:  915  EAPQGNGENKPSENGKVSTGTVENQPTENKPADSLPEAPNEKPVKPENSTDNGMLNPEGN 974
                  +K +  +        +NK  +  P +     KP N         + E
Sbjct:  710  PLKSKKKSDKDQKGSSIKKFFGLKPSDQNKNTKTTQPTPLKSSPKPSNKPVTATVTTEKK 769

Query:  975  VGSDPMLDPALEEAPAVDP---------VQEKLEKFTAS 1004
              +P    A E  P+++P          V  K+E+  AS
Sbjct:  770  ENVEPK-STATETKPSLEPSFSGFSQGSVHNKVEQSDAS 807
```

tr Q879S6 **Hemagglutinin-like secreted protein [pspA] [Xylella**   3457
           fastidiosa   AA
           (strain Temecula1 / ATCC 700964)]   align

 Score = 40.0 bits (92), Expect = 0.23
 Identities = 28/69 (40%), Positives = 33/69 (47%), Gaps = 6/69 (8%)

```
Query:  142  AHADNVRTKDEINRQKQEHVKDNEKVNSNVAVARSQGRYTTNDGYVFNPADIIEDTGNAY 201
             AHA +V T+    R  Q H +N K N     R   Y  +D Y NPADI+EDT    Y
Sbjct: 1594  AHAPDVVTEA---RMYQPHWRKN-KPNGGSGDFRLSSNYDAHDIYYLNPADILEDT--PY 1647

Query:  202  IVPHGGHYH 210
             I P G  H
Sbjct: 1648  ITPDGQKIH 1656
```

tr <u>Q8IJI0</u> **Pre-mRNA splicing factor, putative [PF10_0217] [Plasmodium** 538 AA
        falciparum (isolate 3D7)]

<u>align</u>

 Score = 40.0 bits (92), Expect = 0.23
 Identities = 36/157 (22%), Positives = 72/157 (44%), Gaps = 7/157 (4%)

Query: 807 ENQTDKPSILPQFKRNKAQENSKLDEKVEEPKTSEKVEKEKLSETGN--STSNSTLEEVP 864
            E+++++   L + + N + N+ D K E   +K K+K S + + S   S+ E+
Sbjct: 346 EDRSERRRSLSEDRSNSRKRNASSDFKRELNSDDDKKSKKKRSYSASPGSAYKSSSRELK 405

Query: 865 TVDPVQEKVAKFAESYGMKLENVLFNMDGTIELYLPSGEVIKKNMADFTGEAPQGNGENK 924
            + D   +++++ +SY + +V +N   + E   E +K   G A + G+NK
Sbjct: 406 SQDKSNDRLSESKKSYKSQSASVRYN---STEEKKSDDETVKPKRGG-RGNA-KNTGKNK 460

Query: 925 PSENGKVSTGTVENQPTENKPADSLPEAPNEKPVKPE 961
            ++ K   T EN +++   D   +E+ KPE
Sbjct: 461 NTKGSKKGAKTEENHDSKSVSNDRNKSDKSEEKEKPE 497

tr   <u>O96609</u>   **Surface antigen ariel1 [Entamoeba histolytica]** 215 AA
<u>align</u>

 Score = 40.0 bits (92), Expect = 0.23
 Identities = 42/163 (25%), Positives = 64/163 (38%), Gaps = 18/163 (11%)

Query: 807 ENQTDKPSILPQFKRNKAQENSKLDEKV--EEPKTSEKVEKEKLSETGNSTSNSTLEEVP 864
            EN+ DK   K+ NS+LDE   + ++S   E   N +S++   E
Sbjct: 45 ENEEDK---------KSSSNSELDENSNNQPDESSNNKPNESSDNKPNESSDNKPNESS 94

Query: 865 TVDPVQEKVAKFAESYGMKLENVLFNMDGTIELYLPSGEVIKKNMADFTGEAPQGNGENK 924
               P +   K ES   K   N   P+ E   ++ +   P   + NK
Sbjct: 95 NNKPSESSNNKPDESSNNKPNESSDNKPNESSNNKPN-ESSNNKPSESSNNKPDESSNNK 153

Query: 925 PSENGKVSTGTVENQPTENKPADSLPEAPNEKP-VKPENSTDN 966
            P+E   S+   N+ + NKP +S   PNE   KP +S+DN
Sbjct: 154 PNE----SSDNKPNESSNNKPNESSDNKPNESSNNKPGSSSDN 192

tr <u>O17102</u> **Hypothetical protein F42G2.6 [F42G2.6] [Caenorhabditis** 733
        elegans]   AA
<u>align</u>

 Score = 40.0 bits (92), Expect = 0.23
 Identities = 21/91 (23%), Positives = 45/91 (49%)

Query: 806 KENQTDKPSILPQFKRNKAQENSKLDEKVEEPKTSEKVEKEKLSETGNSTSNSTLEEVPT 865
            KE T KP   + + +K ++N L +   + K +E   EK ++   T N+ E++
Sbjct: 578 KEETTTKPEYRKKLRMSKRKQNELLQKLWLKEKNTENTHAEKTNQEKKLTENAVQEQLEQ 637

Query: 866 VDPVQEKVAKFAESYGMKLENVLFNMDGTIE 896
            + + ++ + A+S +KLE ++ M   ++
Sbjct: 638 MMKMIQEERRAADSKNLKLEQMVLEMKDKLD 668

```
sp_vs  Q28820-    Splice isoform Skeletal 2 of Q28820 [TRDN] [Oryctolagus    689
       5           cuniculus                                                  AA
                   (Rabbit)]                                                  align
```

Score = 40.0 bits (92), Expect = 0.23
Identities = 40/187 (21%), Positives = 69/187 (36%), Gaps = 12/187 (6%)

```
Query:  805  EKENQTDKPSILPQFKRNKAQENSKLDEKVEEPKTSEKVEKEKLSETGNSTSNSTLEEVP 864
                +++ + K + P ++ K ++ K EK EEP  S KE ++        ++ ST + VP
Sbjct:  374  DEKKEDSKKAKKPAEEQPKGKKQEK-KEKHEEPAKSTKKEHAEIERKEEVSAASTKKAVP 432

Query:  865  TVDPVQEKVAKFAESYGMKLENVLFNMDGTIELYLPSGEVIKKNMADFTGEAPQGNGENK 924
                +EK K E    K        G I  L   E+ K+        +   E K
Sbjct:  433  AKK--EEKTTKTVEQETRK------EKPGKISSVLKDKELTKEKEVKVPASLKEKGSETK 484

Query:  925  PSENGKVSTGTVENQPTENKPADSLPEAPNEKPVKPENSTDNGMLNPEGNVGSDPMLDPA 984
                 E        ++   E KP +   P +  +K E   +  ++ PE           +
Sbjct:  485  KDEKTSKPEPQIKK---EEKPGKEVKPKPPQPQIKKEEKPEQDIMKPEKTALHGKPEEKV 541

Query:  985  LEEAPAV 991
             L++  AV
Sbjct:  542  LKQVKAV 548
```

Score = 35.8 bits (81), Expect = 4.4
Identities = 22/82 (26%), Positives = 40/82 (47%), Gaps = 2/82 (2%)

```
Query:  799  VEVPILEKENQTDKPSILPQFKRNKAQENSKLDEKVEEPKTSEKVEKEKLSETGNSTSNS 858
               +   ++ KE + +K  +  + K + + EK E+P+T    ++EK + T
Sbjct:  158  IPTKVVHKEKEKEKEKVKEKEKPEKKATHKEKLEKKEKPETKTVTKEEKKARTKEKIEEK 217

Query:  859  TLEEVPTV--DPVQEKVAKFAE 878
             T +EV  V + V++ VAK  E
Sbjct:  218  TKKEVKGVKQEKVKQTVAKAKE 239
```

```
sp  P19814     Trans-Golgi network integral membrane protein TGN38        357
    TGN3_RAT   precursor [Ttgn1]                                         AA
               [Rattus norvegicus (Rat)]                                  align
```

Score = 39.7 bits (91), Expect = 0.30
Identities = 56/206 (27%), Positives = 84/206 (40%), Gaps = 42/206 (20%)

```
Query:  824  AQENSKLDEKVEEPKTSEKVEKEKLSETGNSTSNSTLEEVPTVDPVQEKVA-KFAESYGM 882
             A+ ++ +  VE    + +K  ET ++    TL+++ VDP QEK  KF +  G
Sbjct:   73  ARTSASVSSGVESATNLNLDDSKKHPETADAKLKETLQQLLPVDPKQEKSGQKFTKDSGS 132

Query:  883  KLENVLFNMDGTIELYLPSGEVIKKNMADFTGEAPQGNGENKP--SENGKVSTGTVENQP 940
                  N  G       G+  K    D   + G    NKP S+N K  TG    N+P
Sbjct:  133  PTGGDSDNTTG-------GDSNKTTGVD--SDKTSGGDSNKPTGSDNDK-PTGGDSNKP 181

Query:  941  TENKPADSLPEAP-------NEKPVKPE--NSTDNG--------MLNPE-GNVGSDPML 981
             T  P+++  E P        EK KP   + T++G          L PE G+  S+P
Sbjct:  182  TSKVPSNT--ETPKIDKVQLTEKGQKPTLISKTESGEKLAGDSDFSLKPEKGDKSSEPTE 239
```

```
Query:  982  DPALEEAPAVD-------PVQEKLEK 1000
             D  +E    D      P++E+ EK
Sbjct:  240  DVETKEIEEGDTEPEEGSPLEEENEK 265
```

sp <u>O51498</u>   **DNA polymerase I (EC 2.7.7.7) (POL I) [polA] [Borrelia**   908
   DPO1_BORBU   **burgdorferi**   AA
                (Lyme disease spirochete)]   <u>align</u>

 Score = 39.7 bits (91), Expect = 0.30
 Identities = 56/274 (20%), Positives = 114/274 (41%), Gaps = 11/274 (4%)

```
Query:  638  NMLVKLITPDGKVLEKVSGKVFGEGVGNIANFELDQPYLPGQTFKYTIASKDYPEVSYDG 697
             N L +I+PD +L+ +S V  + N + E+D Y+ +    +   KDY  + D
Sbjct:  126  NYLTYIISPDKDLLQTMSEYVKILKIENNSFIEMDNEYVTKKFGVNSFQIKDYLAIVGDR 185

Query:  698  TFTVP--TSLAYKMASQTIFYPFHAGDTYLRVNPQFAVPKGTDALVRVFDEFHGNAYLEN 755
             +  +P   + K A+ +    F   D + N+    K  + L++ +    NA+L
Sbjct:  186  SDNIPGIKGIGAKGAA-NLLREFKTLDG-IYSNLEIINKKHRELLIKEKE----NAFLSY 239

Query:  756  NYKVGEIKLPIPKLNQGTTRTAGNKIPVTFMANAYLDNQSTYIVEVPILEKENQTDKPSI 815
             E  L IP++      +  +I  F ++ +    TY ++   EKEN    K
Sbjct:  240  ELVSLEENLKIPEIENFALKNFSEEIISLFEKHSAIALIKTYKKDILKQEKENADQKSLF 299

Query:  816  LPQFKRNKAQENSKLD-EKVEEPKTSEKVEKEKLSETGNSTSNSTLE-EVPTVDPVQEKV 873
             +   N  + + +D E V+   + K+E + L E+     +++ E  ++D    K+
Sbjct:  300  KQEPTTNSLDDINTIDTENVKYRSITTKIELDDLIESLKKAKYISIDTETSSLDTYTAKL 359

Query:  874  AKFAESYGMKLENVLFNMDGTIELYLPSGEVIKK 907
             + S+ + E     ++    ++Y+    +I+K
Sbjct:  360  IGISISF-KEFEGYYIPIEAKGKIYIEKNYIIQK 392
```

sp <u>P42759</u>   **Dehydrin ERD10 (Low-temperature-induced protein LTI45)**   260
   DH10_ARATH   **[ERD10]**   AA
                [Arabidopsis thaliana (Mouse-ear cress)]   <u>align</u>

 Score = 39.7 bits (91), Expect = 0.30
 Identities = 40/162 (24%), Positives = 76/162 (46%), Gaps = 9/162 (5%)

```
Query:  792  DNQSTYIVEVPILEKENQTDKPSILPQFKRNKAQENSKLDEKVEEPKTSEKVEKEKLSET 851
             +++ T + ++    +E + +KPS+L + R+ +  +S DE+ E+ +  +K +K+K+ E
Sbjct:   74  EHKPTLLEQLHQKHEEEEENKPSLLDKLHRSNSSSSSSSDEEGEDGEKKKKEKKKKIVE- 132

Query:  852  GNSTSNSTLEEVPTVDPVQEKVAKFAESYGMKLENVLFNMDG--TIELYLPSGEVIKKNM 909
             G+        E   +D ++EK    E G    V+  M   ++E + P  E  KK
Sbjct:  133  GDHVKTVEEENQGVMDRIKEKF-PLGEKPGGDDVPVVTTMPAPHSVEDHKPEEEE-KKGF 190

Query:  910  ADFTGEAPQGNGENKPSENGKVSTGTVENQPTENKPADSLPE 951
             D   E   G+ + KP ++  V+T +   T   P  +PE
Sbjct:  191  MDKIKEKLPGHSK-KPEDSQVVNTTPLVETAT---PIADIPE 228
```

sp P13952   G2/mitotic-specific cyclin B [Spisula solidissima   428
    CG2B_SPISO (Atlantic   AA
                surf-clam)]   align

 Score = 39.7 bits (91), Expect = 0.30
 Identities = 41/140 (29%), Positives = 58/140 (41%), Gaps = 11/140 (7%)

Query: 837 PKTSEKVEKEKLSETGNSTSNSTLEEVPTVDPVQEKVAKFAESYGMKLENVLFNMDGTIE 896
            P TS ++ L + GN S T+ +VP DP+ +K   S+ K +L   T
Sbjct: 32   PHTSHASQRNTLGDIGNQVSAITISDVPRKDPIIKKEIVHLSSHQHK---ILTKSKATTS 88

Query: 897 L-YLPSGEVIKKNMADFTGEAPQGNGENKPSENGKVSTGTVENQPTENKPAD---SLPEA 952
            L L   I K   FT P   KP+   V T TV QPT   P D   ++PE+
Sbjct: 89   LKSLAEESHIPKKQEAFTFLEPVA-AMPKPT---TVPTATVLPQPTVPVPMDISENVPES 144

Query: 953 PNEKPVKPENSTDNGMLNPE 972
            +   + +N N   NP+
Sbjct: 145 FSRVLLNVQNIDANDKENPQ 164

tr Q63575 Trans golgi network (TGN) specific integral membrane protein 380
            TGN38   AA
            precursor [tgn41] [Rattus norvegicus (Rat)]   align

 Score = 39.7 bits (91), Expect = 0.30
 Identities = 56/206 (27%), Positives = 84/206 (40%), Gaps = 42/206 (20%)

Query: 824 AQENSKLDEKVEEPKTSEKVEKEKLSETGNSTSNSTLEEVPTVDPVQEKVA-KFAESYGM 882
            A+ ++ +   VE   + +K ET ++   TL+++ VDP QEK   KF + G
Sbjct: 73   ARTSASVSSGVESATNLNLDDSKKHPETADAKLKETLQQLLPVDPKQEKSGQKFTKDSGS 132

Query: 883 KLENVLFNMDGTIELYLPSGEVIKKNMADFTGEAPQGNGENKP--SENGKVSTGTVENQP 940
              N G   G+ K   D   + G   NKP S+N K TG   N+P
Sbjct: 133 PTGGDSDNTTG--------GDSNKTTGVD--SDKTSGGDSNKPTGSDNDK-PTGGDSNKP 181

Query: 941 TENKPADSLPEAP-------NEKPVKPE--NSTDNG---------MLNPE-GNVGSDPML 981
            T   P+++ E P   EK KP   + T++G   L PE G+   S+P
Sbjct: 182 TSKVPSNT--ETPKIDKVQLTEKGQKPTLISKTESGEKLAGDSDFSLKPEKGDKSSEPTE 239

Query: 982 DPALEEAPAVD-------PVQEKLEK 1000
            D +E D   P++E+ EK
Sbjct: 240 DVETKEIEEGDTEPEEGSPLEEENEK 265

tr Q9KK42   Surface protein PspC [pspC] [Streptococcus   699 AA
             pneumoniae]   align

 Score = 39.7 bits (91), Expect = 0.30
 Identities = 48/228 (21%), Positives = 94/228 (41%), Gaps = 21/228 (9%)

Query: 786 MANAYLDNQSTYIVEVPILEKENQTDKPS--------ILPFKRN-----KAQENSKLDE 832
            + N D + TY+ E+ +LE +++ + PS   QFK++   ++ ++ +
Sbjct: 89   LINKLQDIKRTYLYELNVLEDKSKAELPSKIKAELDAAFEQFKKDTLPTEPGKKVAEAKK 148

Query: 833 KVEEPKTSEKVEKEKLSETGNSTSNSTLEEVPTVDPVQEKVAKFAESYGMKLENVLFNMD 892

```
                  KVEE +   K +KE+      + +  TLE      V+    K AE  +K E      +
Sbjct: 149  KVEEAEKKAKAQKEEDYRNYPTITYKTLELEIAESDVK---VKEAELELVKKEADESRNE 205

Query: 893  GTIELYLPSGEVIKKNMADFTGEAPQGNGENKPSENGKVSTGTVENQPTENKPADSLPEA 952
                  GTI    +   ++  A+ T          K  E K      E  ++ + +   +
Sbjct: 206  GTIN---QAKAKVESEQAEATRLKKIKTDREKAEEEAKRRADAKEQDESKRRKS-RVKRG 261


Query: 953  PNEKPVKPENSTDNGMLNPEGNVGSDPMLDPALEEAPAVDPVQEKLEK 1000
                  +P  P+    DN  + +  +VG + + P+L+     V  ++K+E+
Sbjct: 262  DLGEPATPDKK-DNDAKSSDSSVGEETLPSPSLKPGKKVADAEKKVEE 308
```

tr Q7P6T9 **TonB protein [FNV1602] [Fusobacterium nucleatum subsp.   254
          vincentii   AA
          ATCC 49256]   align**

```
 Score = 39.7 bits (91), Expect = 0.30
 Identities = 29/77 (37%), Positives = 40/77 (51%), Gaps = 10/77 (12%)

Query: 807  ENQTDKPSILPQFKR--NKAQENSKLDEKV---EEPKTSEKVEKEKLSETGNSTSNSTLE 861
                  E   DKP   P+ +R   KA+EN K ++ V    E+PKT+ K EK  L++       SNS  +
Sbjct: 64   EKPEDKPIEKPEKERPEKKAEENKKAEKTVQVEEKPKTTPKKEKPSLADLKKQISNSQPK 123

Query: 862  E-----VPTVDPVQEKV 873
                        PT DP  E+V
Sbjct: 124  TSNGGFSPTADPDGEEV 140
```

tr Q8H7D1 **Hypothetical protein [Arabidopsis thaliana (Mouse-ear cress)] 346 AA**

                                                                        align

```
 Score = 39.7 bits (91), Expect = 0.30
 Identities = 68/285 (23%), Positives = 107/285 (36%), Gaps = 35/285 (12%)

Query: 714  IFYPFHAGDTYLRVNPQFAVPKGTDA-----LVRVFDEFHGNAYL-ENNYKVGEIKLPIP 767
                  +FYP AG T +    F P G +        L +      GN + E  + GE
Sbjct: 21   LFYPKRAG-TPRKTEIVFVAPTGEEISSRQLEQYLKAHPGNPVISEFEWTTGETPRRSS 79

Query: 768  KLNQGTTRTAGNKIPVTFMANAYL---DNQSTYIVEVPILEKENQ-TDKPSILPQFKRNK 823
                  +++Q    T +K P+    + L   DN+          KEN   DK        +   K
Sbjct: 80   RISQKVKATP-DKEPLLKKRRSSLTKKDNKEAAEKNEEAAVKENMDVDKDGKTENAEAEK 138

Query: 824  AQENSKLDEKVEEPKTS---EKVEKEKLSETGNSTSNSTLEEVPTVDPVQEKVAKFAESY 880
                  +E  + E   E K +   EK EK+++  G T      +    +  +EK + AE+
Sbjct: 139  EKEKEGVTEIAEAEKENNEGEKTEAEKVNKEGEKTEAGKEGQTEIAEAEKEKEGEKAEAE 198

Query: 881  GMKLENVLFNMDGTIELYLPSGEVIKKNMADFTGEAPQGNGENKPSE-------NGKVST 933
                  + E V   D    + + + E+ KK       A  G G  +PS+         K +
Sbjct: 199  NKEAEVV---RDKKESMEVDTSELEKK-------AGSGEGAEEPSKVEGLKDTEMKEAQ 247

Query: 934  GTVENQPTENKPADSLPEAPNEKPVKPENSTDNGMLNPEGNVGSD 978
                  V    E KPA+   E N+  V E + +    E N+ SD
Sbjct: 248  EVVTEADVEKKPAEEKTE--NKGSVTTEANGEQNVTLGEPNLDSD 290
```

tr <u>Q28687</u> **Neurofilament-H (Fragment) [Oryctolagus cuniculus (Rabbit)]** 606 AA

<u>align</u>

Score = 39.7 bits (91), Expect = 0.30
Identities = 44/196 (22%), Positives = 80/196 (40%), Gaps = 16/196 (8%)

```
Query:  817  PQFKRNKAQENSKLDEKVEEPKTSEKVEKEKLSETGNSTSNSTLEEVPTVD----PVQEK  872
             P+  ++  +E +K  EK + P+ ++   KE+ +++      +   EE  + +    PV+E+
Sbjct:  355  PEKAKSPVKEEAKSPEKAKSPEKAKSPVKEE-AKSPEKAKSPVKEEAKSPEKAKSPVKEE  413

Query:  873  VAKFAESYGMKLENV-LFNMDGTIELYLPSGEVIKKNMADFTGEAPQGNGENKPSENGKV  931
              AK  E   K E V +      +E   P+   ++K+      EAP+       E  K
Sbjct:  414  -AKSPEKETPKKEEVKVKEPPKKVEETAPAPPKVEKDSKK--DEAPKKEAPKPAVEKPKE  470

Query:  932  STGTVENQPTENKPADSLPEAPNEKPVKPENSTDNGMLNPEGNVGSDPML-------DPA  984
             ST   +   E+K  + + ++ ++ KP+ T+    PE     +P           P
Sbjct:  471  STAEAKKDEAEDKKKAAPAKMEGKEEAKPKEKTEVAKKEPEDAKAKEPSKPTEKEPEKPK  530

Query:  985  LEEAPAVDPVQEKLEK 1000
             EE PA    +E  E+
Sbjct:  531  KEETPAAPVKKEAKEE 546
```

Database: EXPASY/UniProt
    Posted date: Sep 26, 2004 5:48 PM
  Number of letters in database: 502,311,625
  Number of sequences in database: 1,569,420

Lambda   K   H
  0.309   0.131   0.365

Gapped
Lambda   K   H
  0.267   0.0410   0.140

Matrix: BLOSUM62
Gap Penalties: Existence: 11, Extension: 1
length of query: 1039
length of database: 502,311,625
effective HSP length: 134
effective length of query: 905
effective length of database: 292,009,345
effective search space: 264268457225
effective search space used: 264268457225
T: 11
A: 40
X1: 16 (7.1 bits)
X2: 38 (14.6 bits)
X3: 64 (24.7 bits)
S1: 42 (21.7 bits)
S2: 78 (34.7 bits)

| ExPASy Home page | Site Map | Search ExPASy | Contact us | Proteomics tools | Swiss-Prot |

WEST Search History

Hide Items | Restore | Clear | Cancel

DATE: Monday, September 27, 2004

Hide?	Set Name	Query	Hit Count
		DB=PGPB,USPT,USOC,EPAB,JPAB,DWPI,TDBD; PLUR=YES; OP=AND	
□	L1	(sillery or brodeur or rioux or bizard or charland).in.	828
□	L2	(hamel or pineau or brodeur or rioux or charland).in.	2043
□	L3	L2 and strep$	57

END OF SEARCH HISTORY

First Hit

L3: Entry 1 of 57   File: PGPB   Sep 2, 2004

PGPUB-DOCUMENT-NUMBER: 20040171802
PGPUB-FILING-TYPE: new
DOCUMENT-IDENTIFIER: US 20040171802 A1

TITLE: Haemophilus influenzae antigens and corresponding dna fragments

PUBLICATION-DATE: September 2, 2004

INVENTOR-INFORMATION:

NAME	CITY	STATE	COUNTRY	RULE-47
Hamel, Josee	Sillery		CA	
Couture, France	St.-Ferreol-les-neiges		CA	
Brodeur, Bernard R.	Sillery		CA	
Martin, Denis	St-Augustin-de-Desmaures		FR	
Quellet, Catherine	St-Jean-Chrysostome		CA	
Tremblay, Mireille	Quebec		CA	
Charbonneau, Annie	Quebec		CA	
Vayssier, Catherine	Sillery		CA	

APPL-NO: 10/ 398186   [PALM]
DATE FILED: April 2, 2003

RELATED-US-APPL-DATA:
Application is a non-provisional-of-provisional application 60/236712, filed
October 2, 2000,

PCT-DATA:

DATE-FILED	APPL-NO	PUB-NO	PUB-DATE	371-DATE	102(E)-DATE
Oct 2, 2001	PCT/CA01/01402				

INT-CL: [07] C07 K 14/195, C12 N 1/21, C07 H 21/04

US-CL-PUBLISHED: 530/350; 435/069.3, 435/320.1, 435/252.3, 536/023.7
US-CL-CURRENT: 530/350; 435/252.3, 435/320.1, 435/69.3, 536/23.7

REPRESENTATIVE-FIGURES: NONE

ABSTRACT:

The present invention relates to polypeptides of Haemophilus influenzae which may
be used for prophylaxis, diagnostic and/or therapy purposes.

Generate Collection   Print

Search Results - Record(s) 1 through 50 of 57 returned.

1. 20040171802. 02 Apr 03. 02 Sep 04. Haemophilus influenzae antigens and corresponding dna fragments. Hamel, Josee, et al. 530/350; 435/252.3 435/320.1 435/69.3 536/23.7 C07K014/195 C12N001/21 C07H021/04.

2. 20040171113. 19 Apr 04. 02 Sep 04. Antigens of group b streptococcus and corresponding dna fragments. Martin, Denis, et al. 435/69.1; 435/252.3 435/320.1 530/350 536/23.7 C07H021/04 C12N001/21 C07K014/315.

3. 20040132652. 28 Aug 03. 08 Jul 04. Pharmaceutical compositions. Martin, Denis, et al. 514/12; 424/450 A61K009/127 A61K038/17.

4. 20040122509. 20 Dec 02. 24 Jun 04. Radiopaque ePTFE medical devices. Brodeur, Christopher Brian. 623/1.34; 600/431 623/901 A61F002/06.

5. 20040097706. 18 Nov 03. 20 May 04. Streptococcus pyogenes antigens and corresponding dna fragments. Martin, Denis, et al. 530/350; 424/190.1 C07K014/315 C12Q001/68 A61K039/02.

6. 20040071730. 26 Sep 03. 15 Apr 04. Bvh-a2 and bvh-a3 antigens of group b streptococcus. Martin, Denis, et al. 424/190.1; 435/252.3 435/320.1 435/6 435/69.1 435/7.32 530/350 C07K014/315 C12N015/74 G01N033/554 G01N033/569 C12Q001/68 A61K039/02.

7. 20040029248. 16 Jun 03. 12 Feb 04. CHD5 encoding nucleic acids, polypeptides, antibodies and methods of use thereof. Brodeur, Garrett M., et al. 435/199; 435/320.1 435/325 435/6 435/69.1 536/23.2 C12N009/22 C12Q001/68 C07H021/04 C12P021/02 C12N005/06.

8. 20030232976. 20 Dec 02. 18 Dec 03. Streptococcus antigens. Hamel, Josee, et al. 536/23.1; C07H021/02 C07H021/04.

9. 20030228323. 13 Jan 03. 11 Dec 03. Novel group B streptococcus antigens. Brodeur, Bernard R., et al. 424/184.1; A61K039/00 A61K039/38.

10. 20030198943. 23 Jan 02. 23 Oct 03. Identification of a large number of biological (micro) organisms groups at different levels by their detection on a same array. Remacle, Jose, et al. 435/5; 435/287.2 435/6 435/7.22 435/7.32 C12Q001/70 C12Q001/68 G01N033/53 G01N033/569 G01N033/554 C12M001/34.

11. 20030165528. 22 Apr 03. 04 Sep 03. Streptococcus pyogenes antigens. Martin, Denis, et al. 424/190.1; 435/252.3 435/320.1 435/69.3 530/350 536/23.7 A61K039/02 C07H021/04 C07K014/195 C12P021/02 C12N001/21 C12N015/74.

12. 20030124522. 19 May 00. 03 Jul 03. Method for the identification and/or the quantification of a target compound obtained from a biological sample upon chips. Remacle, Jose, et al. 435/6; 435/7.9 C12Q001/68 G01N033/53 G01N033/542.

13. 20030077293. 20 Jun 01. 24 Apr 03. Streptococcus antigens. Hamel, Josee, et al. 424/190.1; 435/183 435/252.3 435/320.1 435/69.3 536/23.7 A61K039/02 C07H021/04 C12N009/00 C12P021/02

C12N001/21 C12N015/74.

☐ 14. <u>20030059896</u>. 20 Dec 01. 27 Mar 03. Novel chlamydia antigens and corresponding DNA fragments. Couture, France, et al. 435/69.3; 435/252.3 435/320.1 435/7.36 530/350 536/23.7 G01N033/571 C07H021/04 C12P021/02 C12N001/21 C07K014/295 C12N015/74.

☐ 15. <u>20030049271</u>. 21 Feb 02. 13 Mar 03. <u>Streptococcus</u> pyogenes polypeptides and corresponding DNA fragments. Martin, Denis, et al. 424/190.1; 435/252.3 435/320.1 435/69.3 435/7.32 530/350 536/23.7 A61K039/02 G01N033/554 G01N033/569 C07K014/315 C07H021/04 C12P021/02 C12N001/21.

☐ 16. <u>20030031682</u>. 18 Feb 99. 13 Feb 03. NOVEL GROUP B <u>STREPTOCOCCUS</u> ANTIGENS. <u>BRODEUR,</u> BERNARD R., et al. 424/190.1; A61K039/02.

☐ 17. <u>20030007981</u>. 13 Mar 02. 09 Jan 03. Cloning and expression of Haemophilus somnus transferrin-binding proteins. Potter, Andrew A., et al. 424/190.1; 424/256.1 435/252.3 435/320.1 435/69.3 530/350 536/23.7 A61K039/02 A61K039/102 C07H021/04 C07K014/195.

☐ 18. <u>20020106646</u>. 23 Mar 01. 08 Aug 02. Identification of biological (micro) organisms by detection of their homologous nucleotide sequences on arrays. Remacle, Jose, et al. 435/6; 435/287.2 435/91.2 C12Q001/68 C12P019/34 C12M001/34.

☐ 19. <u>6743579</u>. 20 Oct 00; 01 Jun 04. Glycerol as a predictor of glucose tolerance. Gaudet; Daniel, et al. 435/6;. C12Q001/68.

☐ 20. <u>6583333</u>. 12 May 98; 24 Jun 03. Lymphoma-susceptible transgenic mice and methods for studying drug sensitivity of lymphomas. Lowe; Scott W., et al. 800/3; 424/9.1 424/9.2 435/325 435/352 435/354 435/355 800/13 800/14 800/18. G01N033/00 A01K067/033 A01K067/027 C12N005/06 A61K049/00.

☐ 21. <u>6391316</u>. 24 Sep 99; 21 May 02. Vaccine compositions comprising Haemophilus somnus transferrin-binding proteins and methods of use. Potter; Andrew A., et al. 424/256.1; 424/185.1 424/190.1 424/193.1 530/350. A61K039/102.

☐ 22. <u>6287574</u>. 13 Nov 97; 11 Sep 01. Proteinase K resistant surface protein of neisseria meningitidis. <u>Brodeur</u>; Bernard R., et al. 424/250.1; 424/184.1 424/185.1 424/190.1 424/249.1 530/300 530/350 536/23.7. A61K039/095.

☐ 23. <u>6242229</u>. 05 May 98; 05 Jun 01. Cosmetic/pharmaceutical compositions comprising microorganism culture media. <u>Pineau</u>; Nathalie, et al. 435/170; 435/261. A61K035/66.

☐ 24. <u>6100066</u>. 15 Mar 96; 08 Aug 00. Nucleic acid molecules encoding Haemophilus somnus proteins. Potter; Andrew A., et al. 435/69.3; 424/256.1 435/252.3 435/320.1 435/69.1 536/23.1 536/23.7 536/24.32. C12P021/06.

☐ 25. <u>5919620</u>. 07 Jun 95; 06 Jul 99. Heat shock protein HSP72 of <u>Streptococcus</u> pneumoniae. <u>Brodeur</u>; Bernard R., et al. 435/6; 435/4 435/69.1 435/963 536/23.4 536/23.7. C12Q001/06 C12Q001/68 C12P021/06 G01N033/53.

☐ 26. <u>5837204</u>. 18 Apr 97; 17 Nov 98. System for disinfecting the water lines of a dental unit. Prevost; Andre, et al. 422/105; 134/166C 134/169C 137/511 210/136 422/117 422/28 433/80.

A61L002/26 B08B009/06 A61G013/00.

☐ 27. 5780424. 30 Sep 96; 14 Jul 98. Purified ribosomal fractions separated from the nonphotosynthetic filamentous bacteria beggiatoales. Pineau; Nathalie, et al. 514/2; 424/234.1 424/282.1 424/489 424/490 435/820 514/44 514/885 514/886 514/887. A61K038/00 A61K031/70 A61K039/02 A61K009/14.

☐ 28. 5710170. 10 Dec 96; 20 Jan 98. Tri-aryl ethane derivatives as PDE IV inhibitors. Guay; Daniel, et al. 514/332; 514/341 514/365 546/266 546/272.7 548/203. C07D213/30 C07D233/64 A61K031/44 A61K031/415.

☐ 29. 5554530. 04 Aug 94; 10 Sep 96. Aseptic in vitro endomycorrhizal spore mass production. Fortin; J. Andre, et al. 435/256.8; 435/252.5 435/254.1 435/254.11 435/911 47/1.1. C12N001/14 C12N001/04 A01H015/00 A01H017/00.

☐ 30. 5455266. 10 Sep 93; 03 Oct 95. Enhanced chemotherapeutic compositions against microbial infections in fish, the preparation and use thereof. Kusuda; Riichi, et al. 514/450; 514/649 514/653 514/659. A61K031/135 A61K031/335.

☐ 31. 5013664. 28 May 86; 07 May 91. Common protein of Haemophilus influenzae type b identified by a monoclonal antibody. Brodeur; Bernard R., et al. 435/7.32; 435/340 435/70.21 435/810 435/851 435/948 436/548 530/350 530/388.4. C12N005/12 C07K015/78 G01N033/53 G01N033/535.

☐ 32. WO2004007725A1. 15 Jul 03. 22 Jan 04. POLYPEPTIDE OF STREPTOCOCCUS PYOGENES. RIOUX, STEPHANE, et al. C12N015/31; C12N015/62 C12N015/63 C07K014/315 C12N005/10 G01N033/53 A61K038/16.

☐ 33. WO003068813A2. 11 Feb 03. 21 Aug 03. GROUP B STREPTOCOCCUS ANTIGENS. RIOUX, STEPHANE, et al. C07K014/315;.

☐ 34. WO003054007A2. 20 Dec 02. 03 Jul 03. STREPTOCOCCUS ANTIGENS. HAMEL, JOSEE, et al. C07K014/315;.

☐ 35. WO003004650A2. 05 Jul 02. 16 Jan 03. GROUP B STREPTOCOCCUS ANTIGENS AND CORRESPONDING DNA FRAGMENTS. MARTIN, DENIS, et al. C12N015/31; C12N015/63 C07K014/315 C07K019/00 A61K039/09 G01N033/569.

☐ 36. WO002088178A2. 02 May 02. 07 Nov 02. ANTIGENS OF GROUP B STREPTOCOCCUS AND CORRESPONDING DNA FRAGMENTS. MARTIN, DENIS, et al. C07K014/315;.

☐ 37. WO002079475A2. 28 Mar 02. 10 Oct 02. STREPTOCOCCUS PYOGENES ANTIGENS AND CORRESPONDING DNA FRAGMENTS. MARTIN, DENIS, et al. C12N015/31; C12N015/63 C07K014/315 C12N015/62 A61K039/09 C07K016/12 G01N033/53.

☐ 38. WO002066650A2. 21 Feb 02. 29 Aug 02. STREPTOCOCCUS PYOGENES POLYPEPTIDES AND CORRESPONDING DNA FRAGMENTS. MARTIN, DENIS, et al. C12N015/31; C12N015/63 C07K014/315 C07K019/00 A61K039/09 A61P017/00 A61P031/00 G01N033/569.

☐ 39. WO009942588A2. 17 Feb 99. 26 Aug 99. GROUP B STREPTOCOCCUS ANTIGENS. BRODEUR, BERNARD R, et al. C12N015/31; C07K014/315 A61K039/09 C12N001/21.

☐ 40. <u>WO009640928A1</u>. 17 May 96. 19 Dec 96. <u>STREPTOCOCCAL</u> HEAT SHOCK PROTEINS MEMBERS OF THE HSP70 FAMILY. HAMEL, JOSEE, et al. C12N015/31; C07K014/315 C12N001/21 C07K016/12 A61K039/09 A61K039/395 G01N033/569 C12Q001/68.

☐ 41. <u>WO2004048575A</u>. New <u>Streptococcus</u> pneumoniae polynucleotides and polypeptides, useful in preventing or treating otitis media, sinusitis, pneumonia, meningitis, bacteremia, or septicaemia. BLAIS, N, et al. A61K038/16 A61K039/09 C07K014/315 C12N005/10 C12N015/31 C12N015/62 G01N033/53.

☐ 42. <u>WO2004007725A</u>. New polypeptide of <u>Streptococcus</u> pyogenes, useful for the therapeutic or prophylactic treatment of infections, including pharyngitis, erysipelas, impetigo, scarlet fever, bacteremia and necrotizing fascitis. <u>BRODEUR</u>, B R, et al. A61K038/16 C07K014/315 C12N005/10 C12N015/31 C12N015/62 C12N015/63 G01N033/53.

☐ 43. <u>WO2003068813A</u>. New Sip polynucleotide, useful for preparing a composition for diagnosing, treating or preventing <u>Streptococcus</u> infection e.g. sepsis, meningitis, pneumonia, septic arthritis, endocarditis or epiglottis. <u>BRODEUR</u>, B R, et al. C07K014/315.

☐ 44. <u>WO2003054007A</u>. New isolated polypeptides of <u>Streptococcus</u> pneumoniae, useful for diagnosing, preventing or treating <u>streptococcal</u> infection, meningitis, otitis media, bacteremia or pneumonia infection. BLAIS, N, et al. C07H021/02 C07H021/04 C07K014/315.

☐ 45. <u>WO2003004650A</u>. New isolated polypeptide of Group B <u>Streptococcus</u>, useful for diagnosing, preventing or treating <u>streptococcal</u> infections, e.g. sepsis, meningitis, pneumonia, cellulitis, osteomyelitis, septic arthritis, or epiglottis. BOYER, M, et al. A61K039/09 C07K014/315 C07K019/00 C12N015/31 C12N015/63 G01N033/569.

☐ 46. <u>WO 200288178A</u>. New BVH-A4 proteins and genes from serotype III Group B <u>streptococcus</u>, useful for treating or preventing <u>streptococcal</u> infection in infants, pregnant women, non-pregnant adults (e.g. pneumonia), or members of dairy herd (mastitis). BOYER, M, et al. A61K039/09 C07K014/315 C07K019/00 C12N015/31 C12N015/63 G01N033/569 G01N033/68.

☐ 47. <u>WO 200279475A</u>. New isolated polypeptide of <u>Streptococcus</u> pyogenes, useful in manufacturing a medicament for diagnosing, preventing or treating <u>streptococcal</u> infection, pharyngitis, erysipelas, impetigo, scarlet fever, or invasive diseases. <u>BRODEUR</u>, B R, et al. A61K039/09 C07K014/315 C07K016/12 C12N015/31 C12N015/62 C12N015/63 G01N033/53.

☐ 48. <u>FR 2822380A</u>. Preventing adhesion of microbes to skin and mucosa, useful in cosmetics and for treating e.g. acne, using mixture of metal ions, hydrogencarbonate and o-diphenol. <u>PINEAU</u>, N, et al. A61K007/32 A61K007/36 A61K007/48 A61K033/30 A61K033/32 A61K035/78 A61P017/00 A61P017/10 A61P031/10 A61K031:05 A61K031:325 A61K033/32 A61K033:30.

☐ 49. <u>WO 200266650A</u>. New polypeptide useful as a vaccine component for preventing, treating or diagnosing <u>Streptococcus</u> pyogenes infections, e.g. pharyngitis, erysipelas, impetigo, scarlet fever, bacteremia, necrotizing fasciitis or toxic shock. <u>BRODEUR</u>, B R, et al. A61K039/02 A61K039/09 A61P017/00 A61P031/00 C07H021/04 C07K014/315 C07K019/00 C12N001/21 C12N015/31 C12N015/63 C12P021/02 G01N033/554 G01N033/569.

☐ 50. <u>WO 200250107A</u>. Novel polypeptide useful for diagnostic <u>Streptococcal</u> bacterial infection in a host susceptible to <u>Streptococcal</u> infection. <u>HAMEL</u>, J, et al. A61K039/02 A61K039/09 A61P011/04

A61P031/04 A61P039/02 C07K014/195 C07K014/315 C07K019/00 C12N001/15 C12N001/19 C12N001/21 C12N005/10 C12N015/09 C12N015/31 C12N015/63 C12P021/02 C12Q001/04 C12Q001/68 G01N033/53 G01N033/569.

Generate Collection **Print**

Terms	Documents
L2 and strep$	57

Prev Page   Next Page   Go to Doc#

First Hit   Fwd Refs

L3: Entry 25 of 57   File: USPT   Jul 6, 1999

US-PAT-NO: 5919620
DOCUMENT-IDENTIFIER: US 5919620 A

TITLE: Heat shock protein HSP72 of Streptococcus pneumoniae

DATE-ISSUED: July 6, 1999

INVENTOR-INFORMATION:
NAME   CITY   STATE   ZIP CODE   COUNTRY
Brodeur; Bernard R.   Sillery   CA
Martin; Denis   St.-Augustin   CA
Hamel; Josee   Sillery   CA

US-CL-CURRENT: 435/6; 435/4, 435/69.1, 435/963, 536/23.4, 536/23.7

CLAIMS:

    We claim:

    1. An isolated and purified Streptococcus pneumoniae heat shock protein having
    a molecular weight of about 72 kDa as measured by SDS-PAGE.

    2. An isolated and purified protein according to claim 1, comprising amino acid
    acid sequence SEQ ID NO:5.

    3. An isolated and purified protein according to claim 2, consisting of amino
    acid sequence SEQ ID NO:5.

    4. An isolated and purified immunogenic polypeptide comprising an amino acid
    sequence chosen from the group consisting of:

    SEQ ID NO: 5;

    residues 527 to 541 of SEQ ID NO:5 or an immunogenic fragment thereof;

    residues 510 to 607 of SEQ ID NO:5 or an immunogenic fragment thereof;

    residues 439 to 607 of SEQ ID NO:5 or an immunogenic fragment thereof;

    residues 586 to 600 of SEQ ID NO:5 or an immunogenic fragment thereof; and

    residues 244 to 330 of SEQ ID NO:5 or an immunogenic fragment thereof;

    wherein said polypeptide generates a monoclonal antibody not reactive with E.
    coli or S. aureus.

Generate Collection   **Print**

Search Results - Record(s) 1 through 50 of 57 returned.

1. 20040171802. 02 Apr 03. 02 Sep 04. Haemophilus influenzae antigens and corresponding dna fragments. Hamel, Josee, et al. 530/350; 435/252.3 435/320.1 435/69.3 536/23.7 C07K014/195 C12N001/21 C07H021/04.

2. 20040171113. 19 Apr 04. 02 Sep 04. Antigens of group b streptococcus and corresponding dna fragments. Martin, Denis, et al. 435/69.1; 435/252.3 435/320.1 530/350 536/23.7 C07H021/04 C12N001/21 C07K014/315.

3. 20040132652. 28 Aug 03. 08 Jul 04. Pharmaceutical compositions. Martin, Denis, et al. 514/12; 424/450 A61K009/127 A61K038/17.

4. 20040122509. 20 Dec 02. 24 Jun 04. Radiopaque ePTFE medical devices. Brodeur, Christopher Brian. 623/1.34; 600/431 623/901 A61F002/06.

5. 20040097706. 18 Nov 03. 20 May 04. Streptococcus pyogenes antigens and corresponding dna fragments. Martin, Denis, et al. 530/350; 424/190.1 C07K014/315 C12Q001/68 A61K039/02.

6. 20040071730. 26 Sep 03. 15 Apr 04. Bvh-a2 and bvh-a3 antigens of group b streptococcus. Martin, Denis, et al. 424/190.1; 435/252.3 435/320.1 435/6 435/69.1 435/7.32 530/350 C07K014/315 C12N015/74 G01N033/554 G01N033/569 C12Q001/68 A61K039/02.

7. 20040029248. 16 Jun 03. 12 Feb 04. CHD5 encoding nucleic acids, polypeptides, antibodies and methods of use thereof. Brodeur, Garrett M., et al. 435/199; 435/320.1 435/325 435/6 435/69.1 536/23.2 C12N009/22 C12Q001/68 C07H021/04 C12P021/02 C12N005/06.

8. 20030232976. 20 Dec 02. 18 Dec 03. Streptococcus antigens. Hamel, Josee, et al. 536/23.1; C07H021/02 C07H021/04.

9. 20030228323. 13 Jan 03. 11 Dec 03. Novel group B streptococcus antigens. Brodeur, Bernard R., et al. 424/184.1; A61K039/00 A61K039/38.

10. 20030198943. 23 Jan 02. 23 Oct 03. Identification of a large number of biological (micro) organisms groups at different levels by their detection on a same array. Remacle, Jose, et al. 435/5; 435/287.2 435/6 435/7.22 435/7.32 C12Q001/70 C12Q001/68 G01N033/53 G01N033/569 G01N033/554 C12M001/34.

11. 20030165528. 22 Apr 03. 04 Sep 03. Streptococcus pyogenes antigens. Martin, Denis, et al. 424/190.1; 435/252.3 435/320.1 435/69.3 530/350 536/23.7 A61K039/02 C07H021/04 C07K014/195 C12P021/02 C12N001/21 C12N015/74.

12. 20030124522. 19 May 00. 03 Jul 03. Method for the identification and/or the quantification of a target compound obtained from a biological sample upon chips. Remacle, Jose, et al. 435/6; 435/7.9 C12Q001/68 G01N033/53 G01N033/542.

13. 20030077293. 20 Jun 01. 24 Apr 03. Streptococcus antigens. Hamel, Josee, et al. 424/190.1; 435/183 435/252.3 435/320.1 435/69.3 536/23.7 A61K039/02 C07H021/04 C12N009/00 C12P021/02

C12N001/21 C12N015/74.

☐ 14. 20030059896. 20 Dec 01. 27 Mar 03. Novel chlamydia antigens and corresponding DNA fragments. Couture, France, et al. 435/69.3; 435/252.3 435/320.1 435/7.36 530/350 536/23.7 G01N033/571 C07H021/04 C12P021/02 C12N001/21 C07K014/295 C12N015/74.

☐ 15. 20030049271. 21 Feb 02. 13 Mar 03. Streptococcus pyogenes polypeptides and corresponding DNA fragments. Martin, Denis, et al. 424/190.1; 435/252.3 435/320.1 435/69.3 435/7.32 530/350 536/23.7 A61K039/02 G01N033/554 G01N033/569 C07K014/315 C07H021/04 C12P021/02 C12N001/21.

☐ 16. 20030031682. 18 Feb 99. 13 Feb 03. NOVEL GROUP B STREPTOCOCCUS ANTIGENS. BRODEUR, BERNARD R., et al. 424/190.1; A61K039/02.

☐ 17. 20030007981. 13 Mar 02. 09 Jan 03. Cloning and expression of Haemophilus somnus transferrin-binding proteins. Potter, Andrew A., et al. 424/190.1; 424/256.1 435/252.3 435/320.1 435/69.3 530/350 536/23.7 A61K039/02 A61K039/102 C07H021/04 C07K014/195.

☐ 18. 20020106646. 23 Mar 01. 08 Aug 02. Identification of biological (micro) organisms by detection of their homologous nucleotide sequences on arrays. Remacle, Jose, et al. 435/6; 435/287.2 435/91.2 C12Q001/68 C12P019/34 C12M001/34.

☐ 19. 6743579. 20 Oct 00; 01 Jun 04. Glycerol as a predictor of glucose tolerance. Gaudet; Daniel, et al. 435/6;. C12Q001/68.

☐ 20. 6583333. 12 May 98; 24 Jun 03. Lymphoma-susceptible transgenic mice and methods for studying drug sensitivity of lymphomas. Lowe; Scott W., et al. 800/3; 424/9.1 424/9.2 435/325 435/352 435/354 435/355 800/13 800/14 800/18. G01N033/00 A01K067/033 A01K067/027 C12N005/06 A61K049/00.
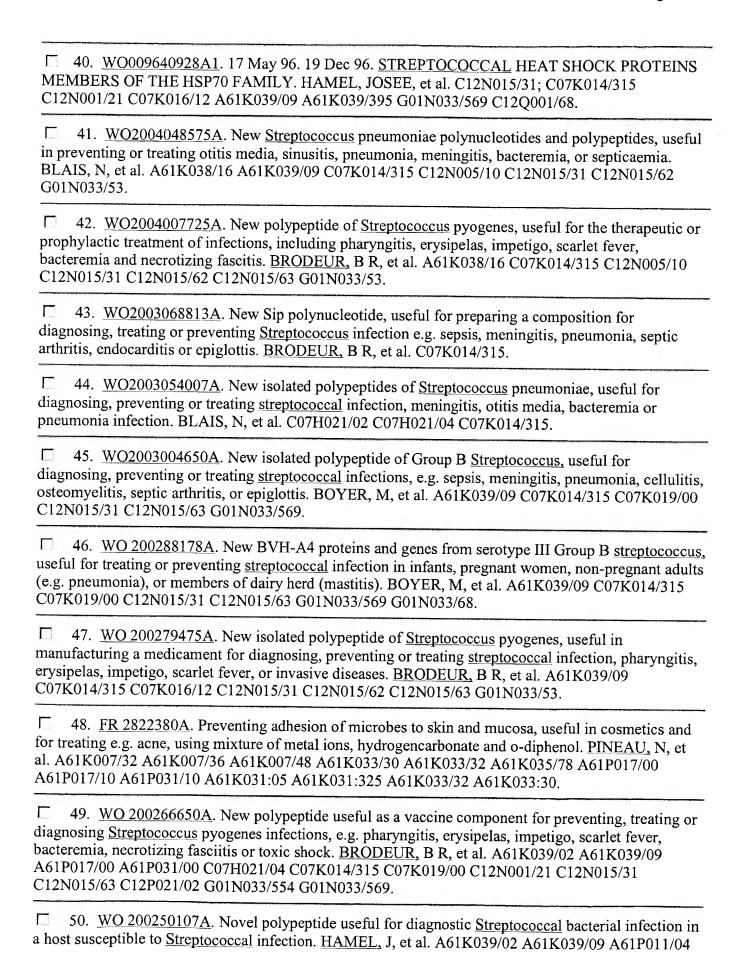

☐ 21. 6391316. 24 Sep 99; 21 May 02. Vaccine compositions comprising Haemophilus somnus transferrin-binding proteins and methods of use. Potter; Andrew A., et al. 424/256.1; 424/185.1 424/190.1 424/193.1 530/350. A61K039/102.

☐ 22. 6287574. 13 Nov 97; 11 Sep 01. Proteinase K resistant surface protein of neisseria meningitidis. Brodeur; Bernard R., et al. 424/250.1; 424/184.1 424/185.1 424/190.1 424/249.1 530/300 530/350 536/23.7. A61K039/095.

☐ 23. 6242229. 05 May 98; 05 Jun 01. Cosmetic/pharmaceutical compositions comprising microorganism culture media. Pineau; Nathalie, et al. 435/170; 435/261. A61K035/66.

☐ 24. 6100066. 15 Mar 96; 08 Aug 00. Nucleic acid molecules encoding Haemophilus somnus proteins. Potter; Andrew A., et al. 435/69.3; 424/256.1 435/252.3 435/320.1 435/69.1 536/23.1 536/23.7 536/24.32. C12P021/06.

☐ 25. 5919620. 07 Jun 95; 06 Jul 99. Heat shock protein HSP72 of Streptococcus pneumoniae. Brodeur; Bernard R., et al. 435/6; 435/4 435/69.1 435/963 536/23.4 536/23.7. C12Q001/06 C12Q001/68 C12P021/06 G01N033/53.

☐ 26. 5837204. 18 Apr 97; 17 Nov 98. System for disinfecting the water lines of a dental unit. Prevost; Andre, et al. 422/105; 134/166C 134/169C 137/511 210/136 422/117 422/28 433/80.

A61L002/26 B08B009/06 A61G013/00.

☐ 27. <u>5780424</u>. 30 Sep 96; 14 Jul 98. Purified ribosomal fractions separated from the nonphotosynthetic filamentous bacteria beggiatoales. <u>Pineau</u>; Nathalie, et al. 514/2; 424/234.1 424/282.1 424/489 424/490 435/820 514/44 514/885 514/886 514/887. A61K038/00 A61K031/70 A61K039/02 A61K009/14.

☐ 28. <u>5710170</u>. 10 Dec 96; 20 Jan 98. Tri-aryl ethane derivatives as PDE IV inhibitors. Guay; Daniel, et al. 514/332; 514/341 514/365 546/266 546/272.7 548/203. C07D213/30 C07D233/64 A61K031/44 A61K031/415.

☐ 29. <u>5554530</u>. 04 Aug 94; 10 Sep 96. Aseptic in vitro endomycorrhizal spore mass production. Fortin; J. Andre, et al. 435/256.8; 435/252.5 435/254.1 435/254.11 435/911 47/1.1. C12N001/14 C12N001/04 A01H015/00 A01H017/00.

☐ 30. <u>5455266</u>. 10 Sep 93; 03 Oct 95. Enhanced chemotherapeutic compositions against microbial infections in fish, the preparation and use thereof. Kusuda; Riichi, et al. 514/450; 514/649 514/653 514/659. A61K031/135 A61K031/335.

☐ 31. <u>5013664</u>. 28 May 86; 07 May 91. Common protein of Haemophilus influenzae type b identified by a monoclonal antibody. <u>Brodeur</u>; Bernard R., et al. 435/7.32; 435/340 435/70.21 435/810 435/851 435/948 436/548 530/350 530/388.4. C12N005/12 C07K015/78 G01N033/53 G01N033/535.

☐ 32. <u>WO2004007725A1</u>. 15 Jul 03. 22 Jan 04. POLYPEPTIDE OF <u>STREPTOCOCCUS</u> PYOGENES. RIOUX, STEPHANE, et al. C12N015/31; C12N015/62 C12N015/63 C07K014/315 C12N005/10 G01N033/53 A61K038/16.

☐ 33. <u>WO003068813A2</u>. 11 Feb 03. 21 Aug 03. GROUP B <u>STREPTOCOCCUS</u> ANTIGENS. RIOUX, STEPHANE, et al. C07K014/315;.

☐ 34. <u>WO003054007A2</u>. 20 Dec 02. 03 Jul 03. <u>STREPTOCOCCUS</u> ANTIGENS. HAMEL, JOSEE, et al. C07K014/315;.

☐ 35. <u>WO003004650A2</u>. 05 Jul 02. 16 Jan 03. GROUP B <u>STREPTOCOCCUS</u> ANTIGENS AND CORRESPONDING DNA FRAGMENTS. MARTIN, DENIS, et al. C12N015/31; C12N015/63 C07K014/315 C07K019/00 A61K039/09 G01N033/569.

☐ 36. <u>WO002088178A2</u>. 02 May 02. 07 Nov 02. ANTIGENS OF GROUP B <u>STREPTOCOCCUS</u> AND CORRESPONDING DNA FRAGMENTS. MARTIN, DENIS, et al. C07K014/315;.

☐ 37. <u>WO002079475A2</u>. 28 Mar 02. 10 Oct 02. <u>STREPTOCOCCUS</u> PYOGENES ANTIGENS AND CORRESPONDING DNA FRAGMENTS. MARTIN, DENIS, et al. C12N015/31; C12N015/63 C07K014/315 C12N015/62 A61K039/09 C07K016/12 G01N033/53.

☐ 38. <u>WO002066650A2</u>. 21 Feb 02. 29 Aug 02. <u>STREPTOCOCCUS PYOGENES</u> POLYPEPTIDES AND CORRESPONDING DNA FRAGMENTS. MARTIN, DENIS, et al. C12N015/31; C12N015/63 C07K014/315 C07K019/00 A61K039/09 A61P017/00 A61P031/00 G01N033/569.

☐ 39. <u>WO009942588A2</u>. 17 Feb 99. 26 Aug 99. GROUP B <u>STREPTOCOCCUS</u> ANTIGENS. BRODEUR, BERNARD R, et al. C12N015/31; C07K014/315 A61K039/09 C12N001/21.

⎡ 40. <u>WO009640928A1</u>. 17 May 96. 19 Dec 96. <u>STREPTOCOCCAL</u> HEAT SHOCK PROTEINS MEMBERS OF THE HSP70 FAMILY. HAMEL, JOSEE, et al. C12N015/31; C07K014/315 C12N001/21 C07K016/12 A61K039/09 A61K039/395 G01N033/569 C12Q001/68.

⎡ 41. <u>WO2004048575A</u>. New <u>Streptococcus</u> pneumoniae polynucleotides and polypeptides, useful in preventing or treating otitis media, sinusitis, pneumonia, meningitis, bacteremia, or septicaemia. BLAIS, N, et al. A61K038/16 A61K039/09 C07K014/315 C12N005/10 C12N015/31 C12N015/62 G01N033/53.

⎡ 42. <u>WO2004007725A</u>. New polypeptide of <u>Streptococcus</u> pyogenes, useful for the therapeutic or prophylactic treatment of infections, including pharyngitis, erysipelas, impetigo, scarlet fever, bacteremia and necrotizing fascitis. <u>BRODEUR, B R</u>, et al. A61K038/16 C07K014/315 C12N005/10 C12N015/31 C12N015/62 C12N015/63 G01N033/53.

⎡ 43. <u>WO2003068813A</u>. New Sip polynucleotide, useful for preparing a composition for diagnosing, treating or preventing <u>Streptococcus</u> infection e.g. sepsis, meningitis, pneumonia, septic arthritis, endocarditis or epiglottis. <u>BRODEUR, B R</u>, et al. C07K014/315.

⎡ 44. <u>WO2003054007A</u>. New isolated polypeptides of <u>Streptococcus</u> pneumoniae, useful for diagnosing, preventing or treating <u>streptococcal</u> infection, meningitis, otitis media, bacteremia or pneumonia infection. BLAIS, N, et al. C07H021/02 C07H021/04 C07K014/315.

⎡ 45. <u>WO2003004650A</u>. New isolated polypeptide of Group B <u>Streptococcus</u>, useful for diagnosing, preventing or treating <u>streptococcal</u> infections, e.g. sepsis, meningitis, pneumonia, cellulitis, osteomyelitis, septic arthritis, or epiglottis. BOYER, M, et al. A61K039/09 C07K014/315 C07K019/00 C12N015/31 C12N015/63 G01N033/569.

⎡ 46. <u>WO 200288178A</u>. New BVH-A4 proteins and genes from serotype III Group B <u>streptococcus</u>, useful for treating or preventing <u>streptococcal</u> infection in infants, pregnant women, non-pregnant adults (e.g. pneumonia), or members of dairy herd (mastitis). BOYER, M, et al. A61K039/09 C07K014/315 C07K019/00 C12N015/31 C12N015/63 G01N033/569 G01N033/68.

⎡ 47. <u>WO 200279475A</u>. New isolated polypeptide of <u>Streptococcus</u> pyogenes, useful in manufacturing a medicament for diagnosing, preventing or treating <u>streptococcal</u> infection, pharyngitis, erysipelas, impetigo, scarlet fever, or invasive diseases. <u>BRODEUR, B R</u>, et al. A61K039/09 C07K014/315 C07K016/12 C12N015/31 C12N015/62 C12N015/63 G01N033/53.

⎡ 48. <u>FR 2822380A</u>. Preventing adhesion of microbes to skin and mucosa, useful in cosmetics and for treating e.g. acne, using mixture of metal ions, hydrogencarbonate and o-diphenol. <u>PINEAU, N</u>, et al. A61K007/32 A61K007/36 A61K007/48 A61K033/30 A61K033/32 A61K035/78 A61P017/00 A61P017/10 A61P031/10 A61K031:05 A61K031:325 A61K033/32 A61K033:30.

⎡ 49. <u>WO 200266650A</u>. New polypeptide useful as a vaccine component for preventing, treating or diagnosing <u>Streptococcus</u> pyogenes infections, e.g. pharyngitis, erysipelas, impetigo, scarlet fever, bacteremia, necrotizing fasciitis or toxic shock. <u>BRODEUR, B R</u>, et al. A61K039/02 A61K039/09 A61P017/00 A61P031/00 C07H021/04 C07K014/315 C07K019/00 C12N001/21 C12N015/31 C12N015/63 C12P021/02 G01N033/554 G01N033/569.

⎡ 50. <u>WO 200250107A</u>. Novel polypeptide useful for diagnostic <u>Streptococcal</u> bacterial infection in a host susceptible to <u>Streptococcal</u> infection. <u>HAMEL, J</u>, et al. A61K039/02 A61K039/09 A61P011/04
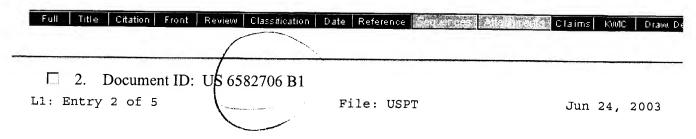
A61P031/04 A61P039/02 C07K014/195 C07K014/315 C07K019/00 C12N001/15 C12N001/19
C12N001/21 C12N005/10 C12N015/09 C12N015/31 C12N015/63 C12P021/02 C12Q001/04
C12Q001/68 G01N033/53 G01N033/569.

Generate Collection   Print

Terms	Documents
L2 and strep$	57

Prev Page   Next Page   Go to Doc#

et al. 1995. "Cloning of an Enterococcus faecalis endocarditis antigen: homology with some adhesins from oral streptococci." Infect. Immun 63:703-706) revealed significant sequence similarity between them. Sequence identities were 81%, 81%, 77%, 82%, and 57%, respectively, for PsaA (S. pneumoniae strain R36A), SsaB (S. sanguis), FimA (S. parasanguis), ScaA (S. gordonii) and EfaA (E. faecalis). Additionally, all six sequences showed great similarity in organization. They have a a hydrophobic leader peptide containing the prolipoprotein consensus sequence LXXC (for signal peptidase II cleavage) within the first 17-20 amino acids. This N-terminal leader sequence appears to represent the area of greatest variability. It is followed by a region of high similarity from amino acids 36-150. The region from 150 to 198 is a variable region and is followed by another conserved region (198-309).

| Full | Title | Citation | Front | Review | Classification | Date | Reference | | | Claims | KWIC | Draw D |

2. Document ID: US 6582706 B1

L1: Entry 2 of 5   File: USPT   Jun 24, 2003

DOCUMENT-IDENTIFIER: US 6582706 B1
TITLE: Vaccine compositions comprising Streptococcus pneumoniae polypeptides having selected structural MOTIFS

Detailed Description Text (65):
The genomic DNA used as target for amplification was isolated from S. pneumoniae Norway strain (serotype 4), the same strain used for genomic sequencing. The complete sequence of the Sp36 gene (SEQ ID NO:9), and its predicted amino acid sequence (SEQ ID NO:8), are given in the Sequence Listing appended hereto. It was noted that the predicted amino acid sequence included a hydrophobic leader sequence followed by a sequence (LSVC) similar to the consensus sequence for Type II signal peptidase (LxxC, in which both x's typically represent small amino acids). Primers (listed as SEQ ID NOS:1-3) were designed that would amplify the Sp36 gene and allow its cloning into pQE10 and expression as a histidine-tagged protein lacking the signal sequence for purification by nickel-affinity chromatography. Cloning of the fragment amplified by SEQ ID Nos 1 and 3 would result in a protein containing amino acids 21 through 819 of Sp36; cloning of the fragment amplified by SEQ ID Nos 2 and 3 would result in a protein containing amino acids 26 through 819 of Sp36 (amino acid numbers refer to SEQ ID NO:8).

| Full | Title | Citation | Front | Review | Classification | Date | Reference | | | Claims | KWIC | Draw D |

3. Document ID: US 6406883 B1

L1: Entry 3 of 5   File: USPT   Jun 18, 2002

DOCUMENT-IDENTIFIER: US 6406883 B1
TITLE: Lmb gene of Streptococcus agalactiae

Detailed Description Text (25):

☐   　Generate Collection　　 　Print

L2: Entry 2 of 24   File: PGPB   Apr 29, 2004

DOCUMENT-IDENTIFIER: US 20040081662 A1
TITLE: Vaccine

Summary of Invention Paragraph:
[0010] In one aspect, the present invention is an immunogenic composition
comprising at least 2 S. pneumoniae proteins selected from the group consisting of
Poly Histidine Triad family (PhtX), Choline Binding Protein family (CbpX), CbpX
truncates, LytX family, LytX truncates, CbpX truncate-LytX truncate chimeric
proteins, pneumolysin (Ply), PspA, PsaA, Sp128, Sp101, Sp130, Sp125 and Sp133. In a
preferred embodiment, one of the proteins is from the Poly Histidine Triad family
(PhtX). In another preferred embodiment, one of the proteins is from the Choline
Binding Protein family (CbpX), or CbpX truncates, or CbpX truncate-LytX truncate
chimeric proteins.

Summary of Invention Paragraph:
[0016] The Streptococcus pneumoniae proteins of the invention are either surface
exposed, at least during part of the life cycle of the pneumococcus, or are
proteins which are secreted or released by the pneumococcus. Preferably the
combination of proteins of the invention are selected from 2 different categories
such as proteins having a Type II Signal sequence motif of LXXC (where X is any
amino acid, e.g., the polyhistidine triad family (phtX)), choline binding proteins
(CbpX), proteins having a Type I Signal sequence motif (e.g., Sp101), proteins
having a LPXTG motif (where X is any amino acid, e.g., Sp128, Sp130), toxins (e.g.,
Ply), etc. Preferred examples within these categories (or motifs) are the following
proteins, or immunologically functional equivalents thereof.

Summary of Invention Paragraph:
[0017] The immunogenic composition of the invention comprises at least 2 proteins
selected from the group consisting of the Poly Histidine Triad family (PhtX),
Choline Binding Protein family (CbpX), CbpX truncates, LytX family, LytX truncates,
CbpX truncate-LytX truncate chimeric proteins (or fusions), pneumolysin (Ply),
PspA, PsaA, Sp128, Sp101, Sp130, Sp125 and Sp133. However, if CbpX is PspC, then
the second protein is not PspA or PsaA. Preferably, the immunogenic composition
comprises 2 or more proteins selected from the group consisting of the Poly
Histidine Triad family (PhtX), Choline Binding Protein family (CbpX), CbpX
truncates, LytX family, LytX truncates, CbpX truncate-LytX truncate chimeric
proteins (or fusions), pneumolysin (Ply), PspA, PsaA, and Sp128. More preferably,
the immunogenic composition comprises 2 or more proteins selected from the group
consisting of the Poly Histidine Triad family (PhtX), Choline Binding Protein
family (CbpX), CbpX truncates, LytX family, LytX truncates, CbpX truncate-LytX
truncate chimeric proteins (or fusions), pneumolysin (Ply), and Sp128

Summary of Invention Paragraph:
[0019] With regards to the PhtX proteins, PhtA is disclosed in WO 98/18930, and is
also referred to Sp36. As noted above, it is a protein from the polyhistidine triad
family and has the type II signal motif of LXXC.

Summary of Invention Paragraph:
[0020] PhtD is disclosed in WO 00/37105, and is also referred to Sp036D. As noted

above, it also is a protein from the polyhistidine triad family and has the type I
LXXC signal motif.

Summary of Invention Paragraph:
[0021] PhtB is disclosed in WO 00/37105, and is also referred to Sp036B. Another
member of the PhtB family is the. C3-Degrading Polypeptide, as disclosed in WO
00/17370. This protein also is from the polyhistidine triad family and has the type
II LXXC signal motif. A preferred immunologically functional equivalent is the
protein Sp42 disclosed in WO 98/18930. A PhtB truncate (approximately 79 kD) is
disclosed in WO99/15675 which is also considered a member of the PhtX family.

CLAIMS:

1. An immunogenic composition comprising at least 2 S. pnumoniae proteins wherein
one of the proteins is selected from the polyhistidine triad family (PhtX) and
another protein is selected from the group consisting of Choline Binding Protein
family (CbpX), CbpX truncates, LyX family, LytX truncates, CbpX truncate-LytX
truncate chimeric proteins, pneumolysin (Ply), PspA, PsaA, Sp128, Sp101, Sp130,
Sp125 and Sp133.

2. An immunogenic composition comprising at least 2 S. pneumoniae proteins wherein
one of the proteins is selected from the group consisting of Choline Binding
Protein family (CbpX), CbpX truncates, and CbpX truncate-LytX truncate chimeric
proteins and another protein selected from the group consisting of polyhistidine
triad family (PhtX), LytC, pneumolysin (Ply), PsaA, and Sp128.